TESIS DOCTORAL

VNIVERSITAT DE VALÈNCIA
INSTITUTO DE CIENCIA MOLECULAR

Theoretical study of the molecular bases that control photochemical processes with biological and nanotechnological interest

TESIS DOCTORAL PRESENTADA POR

Vicenta Saurí Peris

Dirigida por
Dra. Manuela Merchán
Dr. Josep M. Oliva

Valencia 2013

Programa de doctorado en Química Teórica
y Modelización Computacional

Me lo contaron y lo olvidé,

lo vi y lo entendí,

lo hice y lo aprendí.

Confucio (551 a.C. - 479 a.C.)

Dña. Manuela Merchán, Catedrática del Departamento de Química Física y del Instituto de Ciencia Molecular de la Universitat de València, y D. Josep María Oliva, Investigador del Consejo Superior de Investigaciones Científicas en el Instituto Rocasolano de Madrid, certifican que la memoria presentada por la Licenciada Vicenta Saurí Peris con el título **Theoretical study of the molecular bases that control photochemical processes with biological and nanotechnological interest** corresponde a su Tesis Doctoral y ha sido realizada bajo su dirección, autorizando mediante este escrito la presentación de la misma.

En Valencia, Marzo de 2013

Dra. Manuela Merchán Dr. Josep M. Oliva

Doctoranda: Vicenta Saurí Peris

A mi familia

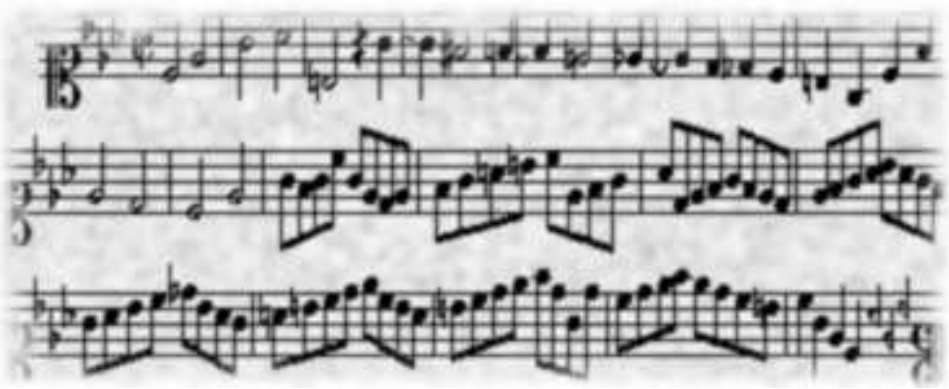
Agradecimientos

Hace cinco años, mi interés se centraba en cursar un máster en ciencia básica. Contaba con dos opciones, un máster en física aplicada y el máster en química teórica y modelización computacional. Después de hablar con el profesor Benito Gimeno, un excelente docente e investigador en física de la Universitat de València, él me animó a matricularme en un máster de óptica, sin embargo, me decliné por la Química. En el curso M1 del máster conocí a la persona que me inició en el mundo de la ciencia, el profesor Luis Serrano Andrés (1966-2010), al que agradezco su calidad humana y científica. Mi aportación en los trabajos de la presente tesis se la dedico a él. También agradezco a mis actuales directores de tesis, los profesores Manuela Merchán y Josep María Oliva, su entrega en sacar adelante esta tesis y su excelente supervisión. Además, me gustaría señalar el trabajo de los doctores Steven Vancoillie y Daniel Roca Sanjuán y de los profesores Kristin Pierloot y Roland Lindh. También quería agradecer a la doctora Remedios González Luque sus comentarios y sugerencias en este trabajo, al profesor José Sánchez Marín por su disponibilidad y ayuda en los comienzos de esta etapa, a María Jesús Leiva por sus gestiones administrativas y a Martín Córdoba y Alejandro Gómez Merchán por proporcionarme el material necesario para el desarrollo de mi trabajo. Tampoco quiero olvidarme de Vicente Pérez Mondéjar y Alejandro Soriano que tantas dudas me han resuelto. Y esto hace que recuerde a Dirac, Roos, Multivac, Qcexval y Bach, las máquinas que me han facilitado el trabajo. Quiero expresar mi gratitud a la Universitat de València y al Ministerio de Educación y Ciencia, y en especial

al grupo QCEXVAL, por darme la oportunidad de vivir esta experiencia. Y por último, doy las gracias a mis compañeros y todos los miembros del grupo QCEXVAL por los buenos momentos que hemos pasado juntos.

Durante esta etapa, he vivido el adiós de mi abuela y la venida de mi sobrina, Irene. Desde estas líneas quiero decir a mis hermanas que siempre han sido un ejemplo a seguir para mí. Mi hermano es una de las personas que más ha influido en mí, por su insaciable interés por el saber, por su facilidad de asombro, por su ilusión. Y nunca podré devolver a mis padres todo el amor que me han regalado en esta treintena de años. Es por ello que dedico esta tesis a mi familia, en especial a mis padres, Carmen y Vicente.

Y quería comenzar la lectura de esta memoria de Tesis Doctoral con un poco de música, mi gran compañera desde mis diez años.



Crab Canon composed by Johann Sebastian Bach (1685-1750). This is the first Canon from Bach's *Musical Offering*. It was written to be played backwards and forwards at the same time. Recently, graphic artist Jos Leys has shown that this piece is basically the musical version of a Möbius strip.

Preface

The present work is framed in the research lines of the QCEXVAL (Quantum Chemistry of the Excited State University of Valencia) group, which was founded in the early 1993 in the Theoretical Chemistry group of the Physical Chemistry Department of the University of Valencia. Mainly, this group is internationally known by its theoretical studies on the photochemistry of molecular systems with quantum-chemical methods, which are introduced in the present thesis. The main goals of the group include the theoretical characterization of electronic transitions and properties of the excited states, assignment of spectra, description of decay pathways, determination of structures, and the study of photoreactivity of neutral or charged systems, especially of biological interest. In the last years, reported works of major impact are related to the photochemistry of DNA. Moreover, studies of molecular systems with applications on photodynamic therapy and nanotechnology are also interesting research subjects for the group.

The development of quantum-chemical methods is important to treat new systems that were not addressed before or to obtain more accurate theoretical results on the molecular properties. The RASPT2 (Multiconfigurational Second-Order Perturbation Theory Restricted Active Space) method, implemented in the MOLCAS software, has been tested in the present thesis to design strategies for computation of organic and inorganic systems. The ad-

vantage of the method is to be accurate in the calculation of excited states for small and medium-sized systems with a reasonable computational cost.

RASPT2 can also be used to compute potential energy surfaces in the excited state, as it has been done in the ethene molecule as a function of the carbon-carbon distance. Nevertheless, the calculation of excited states is not determined by a unique methodology. In fact, in this work, the Hubbard model is parameterized to compute structures and excitation energies of conjugated and aromatic hydrocarbons.

The study in the excited state of systems with applications on nanotechnology has also been addressed in the present thesis. In particular, the interest is focused on boron hydrides which can be used as dopant materials in p-type silicon semiconductors. Moreover, derivatives of these boranes can be applied in photodynamic therapy, since the formation of singlet oxygen is of rising importance and, as being highly reactive, the singlet oxygen can be harmful against cancerous cells.

Another present challenge is to know, precisely, the origin of mutations in DNA. This thesis also treats the study of tautomerization in the excited state of the Watson-Crick (WC) base pairs, via proton or hydrogen transfer. The findings explain the spontaneous mutations of Adenine-Thymine and Guanine-Cytosine, which were predicted by Per-Olov Löwdin in the early sixties. This study has been complemented taking into account the effects of the biological environment with hybrid methods QM/MM which are explained in the dissertation of the thesis. In addition, dynamical calculations have been performed to determine which hydrogen/proton transfer process dominates in the photostability and tautomerism of the WC base pairs.

List of Publications

V. Sauri, L. Serrano-Andrés, A.R.M. Shahi, L. Gagliardi, S. Vancoillie, and K. Pierloot. Multiconfigurational second-order perturbation theory restricted active space (RASPT2) method for electronic excited states: A benchmark study. *Journal of Chemical Theory and Computation*, 7(1):153-168, **2011**.

T.G. Schmalz, L. Serrano-Andrés, V. Sauri, M. Merchán, and J.M. Oliva. A distance-dependent parameterization of the extended Hubbard model for conjugated and aromatic hydrocarbons derived from stretched ethene. *The Journal of Chemical Physics*, 135(19):194103, **2011**.

M.G.S. Londesborough, D. Hnyk, J. Bould, L. Serrano-Andrés, V. Sauri, J.M. Oliva, P. Kubát, T. Polívka, and K. Lang. Distinct photophysics of the isomers of $B_{18}H_{22}$ explained. *Inorganic Chemistry*, 51(3):1471-1479, **2012**.

J.P. Gobbo, V. Sauri, D. Roca-Sanjuán, L. Serrano-Andrés, M. Merchán, and A.C. Borin. On the deactivation mechanisms of the Adenine-Thymine Base Pair. *The Journal of Physical Chemistry B*, 116(13):4089-4097, **2012**.

V. Sauri, J.P. Gobbo, J.J. Serrano-Pérez, M. Lundberg, P.B. Coto, L. Serrano-Andrés, A.C. Borin, R. Lindh, M. Merchán, and D. Roca-Sanjuán. Proton/Hydrogen transfer mechanisms in the Guanine-Cytosine base pair: Photostability and Tautomerism. *Journal of Chemical Theory and Computation*, 9(1):481-496, **2013**.

Contents

Contents 13

1 Introduction 17

1.1 The RASPT2 method 17

1.2 The Hubbard model 32

1.3 Boron Hydrides 42

1.4 Photochemistry of the DNA base pairs 50

2 Objectives 59

3 Quantum-Chemical Methodology 61

3.1 The Hartree-Fock Method 64

3.2 Configuration Interaction (CI) 67

3.3 Many-Body Perturbation Theory (MBPT) 69

3.4 Coupled Cluster (CC) 70

3.5 Hybrid Methods (QM/MM) 71

3.6 Nonadiabatic dynamics by the SH method 75

4 Basis Sets 81

5 Photochemical and Photophysical concepts 93

6 Results 113

6.1 The RASPT2 method for Electronic Excited States. 113

6.1.1 Free Base Porphin 117

6.1.2 Ethene, Benzene and the Valence-Rydberg mixing problem 123

6.1.3 Transition metal compounds and the double *d*-shell effect: the nickel atom and the copper tetrachloride dianion 130

6.2 The Hubbard model for Conjugated and Aromatic Hydrocarbons. 137

6.2.1 Computation of potential energy curves for stretched ethene . . . 138

6.2.2 Fitting of the model parameters 142

6.2.3 Testing and validation of the model 147

6.3 Photophysics of the $B_{18}H_{22}$ boron hydrides 155

6.3.1 Experimental procedure and theoretical methodology 156

6.3.2 Experimental Results 158

6.3.3 Theoretical calculations and comparison with experimental data 164

6.4 Proton/Hydrogen transfer processes in the AT and GC base pairs 179

6.4.1 Mechanisms of photostability and tautomers production in the isolated AT and GC base pairs 190

6.4.2 GC base pair in the biological environment 204

6.4.3 Excited-state dynamics of the GC base pair 207

7 Conclusions 213

8 Resumen 221

List of Terms 239

List of Tables 245

List of Figures 249

Bibliography 257

Chapter 1

Introduction

1.1. The Multiconfigurational Second-Order Perturbation Theory Restricted Active Space (RASPT2) Method

The most awkward restriction of the Complete Active Space Self-Consistent Field (CASSCF) method [1] is the limited number of symmetry- and spin-adapted Configuration State Functions (CSFs) comprising the multireference space, that is, the limited number of electrons and orbitals which constitute the selected active space. The method can be defined as open-ended: as more orbitals are included in the calculation, more CSFs are in general generated and more extensive is the computational process.

To enlarge the active space and therefore the applicability of the methodology, one possibility is to introduce restrictions in the size of the Configuration Interaction (CI) expansion of the wave function. The procedure is known now for some years and was coined as the RASSCF (Restricted Active Space Self-

Consistent Field) method [2, 3]. The active space is in this case subdivided into three subspaces: RAS1 comprising doubly occupied orbitals, RAS2 for orbitals whose occupation numbers take a value between two and zero, and RAS3 for unoccupied orbitals. Other more generalized schemes have been also applied [4]. RAS2 orbitals are treated like in CASSCF calculations, which means, a full CI is applied over them. The configurations obtained in this process should be the primary ones which are essential for describing the non-dynamical correlation effects. On the other hand, the orbitals in RAS1 and RAS3 are treated by excited configurations: a maximum number of excitations, mostly up to four, may occur from RAS1 into the other two subspaces and, also, a maximum number of excitations, up to four, may take place into RAS3 from RAS1 and RAS2 (see Figure 1.1). The variational principle is used to optimize both CI coefficients and molecular orbitals of the RAS wave function.

Provided that the maximum number of excited configurations in RAS1 and RAS3 is limited, the number of CSFs generated in the RAS model can be smaller than in the CASSCF method over the same active space. This feature allows multiconfigurational calculations with larger active spaces and therefore extending the applicability of the multiconfigurational approaches to a large amount of new systems and chemical problems. From now on CASSCF is regarded as a special case of RASSCF.

Nevertheless, the RASSCF method is not as solid as CASSCF. Its convergence is not as well-behaved as in CASSCF and it is maybe even more dependent to the quality of the starting orbitals used in the calculation. Therefore, besides making a good choice of the active space, it is also important to obtain proper starting orbitals before running the suitable calculation, something that is applied also to its methodological predecessor. Up to recently, as dynamical correlation was not available on top of the RAS wave function, the utility of the RASSCF approach was limited to perform geometry optimizations or educated selections of the active space based on the occupancy of the natural orbitals for CASSCF calculations. The addition of the second order perturbation treatment to the RASSCF formulation originating the RASPT2 method made possible to enhance the applicability of the approach to quantitative

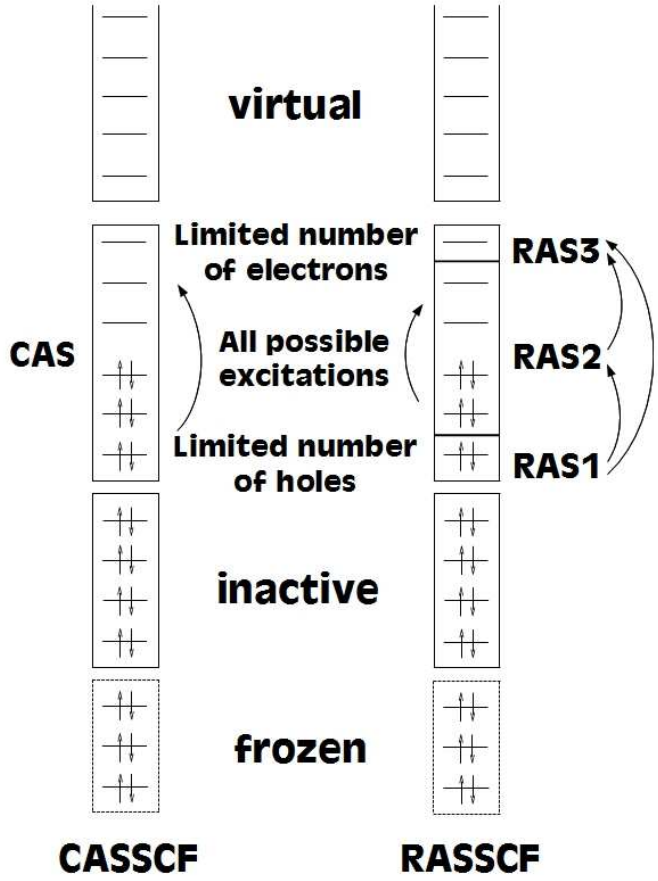


Figure 1.1: Partition of the CAS and RAS orbital subspaces. Taken from Ref. [5]

chemical problems [6, 7]. An initial, hardly employed, implementation of the RASPT2 method was reported earlier [8].

The RASPT2 model extends the many-body perturbation theory (MBPT) up to second order of accuracy. MBPT has the feature of being size-extensive because it scales correctly with the number of particles in the system. The CASPT2 approach [9], in particular its MOLCAS implementation [7], has been proved to be almost size-extensive, unlike truncated-CI approaches like multireference CI. Because of the restrictions in the excitation level in RAS1 and RAS3, the size-extensivity problems may increase in RASPT2, unless fully internal excitations are actually included in a more complex formulation that the one developed so far [6].

As any post-Hartee-Fock (HF) method, the RASPT2 approach focuses on the accurate inclusion of the correlation energy ($E_{corr} = \varepsilon_0 - E_0$), defined as the difference between the exact electronic energy (ε_0) and the Hartee-Fock energy in the limit of a complete one-electron basis set (E_0). The multiconfigurational methods such as CASSCF define a partition of the correlation energy into two: the static, short-range or non-dynamical correlation, basically that originated from near-degenerated configurations strongly changing during the chemical process, and the long-range or dynamical correlation, the remaining one. This is actually an arbitrary partition in which the CSFs included in the CAS CI treatment define the non-dynamical correlation, whereas the correlation energy recovered by the CASPT2 method is named as dynamical correlation. Only by inclusion of both types of correlation the chemical problem can be treated quantitatively, because the effects of the dynamical correlation are clearly differential in many cases for the different states. For the RASSCF/RASPT2 approaches the same type of reasoning applies, although in this case the partition limits might be somewhat diffuse. Normally, certain ammount of dynamical correlation is already taken into account at the RASSCF level.

The formulation of the RASPT2 method is similar to that of the CASPT2 approach, and they can be considered together. In MBPT, the Hamiltonian can be partitioned in some way into a zeroth order term and a perturbation where its strength is governed by a parameter λ:

$$\hat{H} = \hat{H}_0 + \lambda \hat{H}_1 \tag{1.1}$$

The wave function and energy are expanded in a similar way and it is assumed that the zeroth order term Ψ_0 is an eigenfunction of $\hat{H}_0$ with the eigenvalue E_0:

$$|\Psi\rangle = |\Psi_0\rangle + \lambda|\Psi_1\rangle + \lambda^2|\Psi_2\rangle + ... \tag{1.2}$$

$$E = E_0 + \lambda E_1 + \lambda^2 E_2 + ... \tag{1.3}$$

Let us insert 1.2 and 1.3 into the Schrödinger equation and collect terms of the same order in λ:

$$\begin{aligned} \hat{H}_0|\Psi_0\rangle &= E_0|\Psi_0\rangle \\ \left(\hat{H}_0 - E_0\right)|\Psi_1\rangle &= \left(E_1 - \hat{H}_1\right)|\Psi_0\rangle \\ \left(\hat{H}_0 - E_0\right)|\Psi_2\rangle &= \left(E_1 - \hat{H}_1\right)|\Psi_1\rangle + E_2|\Psi_0\rangle \end{aligned} \tag{1.4}$$

It is assumed that the perturbed wave functions are orthogonal to the zeroth order function, $\langle\Psi_0|\Psi_i\rangle = \delta_{0i}$ and this gives intermediate normalization of the total wave function $\langle\Psi|\Psi_0\rangle = 1$. Using this normalization we obtain the following expressions for the energies up to second order:

$$\begin{aligned} E_0 &= \langle \Psi_0 | \hat{H}_0 | \Psi_0 \rangle \\ E_1 &= \langle \Psi_0 | \hat{H}_1 | \Psi_0 \rangle \\ E_2 &= \langle \Psi_0 | \hat{H}_1 | \Psi_1 \rangle \end{aligned} \tag{1.5}$$

The first order wave function appears in the expression for the second order energy. It is obtained by solving the second of the equations in 1.4. It is chosen $|\Psi_0\rangle$ as a reference CAS wave function (the orbitals need not to be optimized but the CAS CI space will be assumed diagonal). The zeroth order Hamiltonian can be expressed as:

$$\hat{H}_0 = \hat{P}_0 \hat{F} \hat{P}_0 + \hat{P}_K \hat{F} \hat{P}_K + \hat{P}_{SD} \hat{F} \hat{P}_{SD} + \hat{P}_X \hat{F} \hat{P}_X \tag{1.6}$$

where the CI space has been partitioned into four subspaces: V_0 (one-dimensional space spanned by the CAS reference function for the state under consideration), V_K (the space spanned by the orthogonal complement to the reference wave function in the restricted full CI subspace used to generate the CAS wave function), V_{SD} (space spanned by all singly and doubly excited CSFs with respect to the CAS reference) and V_X (the rest of the CI space). $\hat{P}_0, \hat{P}_K, \hat{P}_{SD}$ and $\hat{P}_X$ are the corresponding projector operators for the different configurations spaces according to the label. $\hat{F}$ is defined such that for inactive and external orbitals their diagonal elements correspond to orbital energies in the sense of the Koopmans' theorem. $\hat{F}$ can be written as the one-electron operator:

$$\hat{F} = \sum_{p,q} f_{pq} \hat{E}_{pq} \tag{1.7}$$

$$f_{pq} = h_{pq} + \sum_{r,s} D_{rs} \left[(pq|rs) - \frac{1}{2}(pr|qs) \right] \tag{1.8}$$

where $\hat{E}_{pq}$ is the excitation operator used in spin independent operators, h_{pq} is the one-electron Hamiltonian and D_{rs} is the first order density matrix.

The CASSCF wave function is invariant against rotations among the inactive orbitals, the active orbitals, and among the external orbitals. This is used to simplify the matrix f. Let us choose the three subsets of orbitals such that the matrix is diagonal within each subset.

$$\begin{aligned}\hat{F} = &\sum_i \epsilon_i \hat{E}_{ii} + \sum_t \epsilon_t \hat{E}_{tt} + \sum_a \epsilon_a \hat{E}_{aa} + \sum_{i,t} f_{ti} \left[\hat{E}_{it} + \hat{E}_{ti}\right] + \\ &+ \sum_{i,a} f_{ai} \left[\hat{E}_{ia} + \hat{E}_{ai}\right] + \sum_{t,a} f_{at} \left[\hat{E}_{ta} + \hat{E}_{at}\right]\end{aligned} \tag{1.9}$$

where i labels inactive, t active and a external orbitals and it has been assumed that f is a symmetric matrix. For a converged CASSCF reference all f_{ai} matrix elements are zero, but it is not used this simplification because we want this formalism also works for CAS wave functions with unoptimized orbitals. Only those configurations which interact directly with the CAS reference function have to be included into the first order wave function. They all belong to the SD space. Thus, the first order wave function expanded in configurations, $|\Psi_1\rangle = \sum_\mu c_\mu \Phi_\mu$, can be written as:

$$\begin{aligned}|\Psi_1\rangle &= \sum_{p,q,r,s} c_{pqrs} |pqrs\rangle \\ |pqrs\rangle &= \hat{E}_{pq}\hat{E}_{rs}|\Psi_0\rangle\end{aligned} \tag{1.10}$$

where all four indices p, q, r, s cannot be active, since the generated function then belongs to the CAS CI space V_K which do not interact with $|\Psi_0\rangle$. The function space defined in 1.10 (V_{SD}) will cover the entire interacting space

that contributes to the first order wave function. Each term in this expansion, $|pqrs\rangle$, consists of many CSFs. They are linear combinations of CSFs with coefficients from the CAS reference function $|\Psi_0\rangle$. All singly and doubly excited states are included in the wave function except those which have all four indices in the active space. The different functions that contribute to the SD space are, however, in general not orthogonal to each other or even not always linearly independent. The first order equation for the coefficients c_{pqrs} in 1.10 takes the more general form:

$$(\mathbf{F} - E_0\mathbf{S})\,\mathbf{C} = -\mathbf{V} \tag{1.11}$$

where $\mathbf{F}$ is the Fock matrix in the space V_{SD}, $\mathbf{S}$ is the corresponding overlap matrix, $\mathbf{C}$ is the vector of the coefficients and $\mathbf{V}$ is the vector representing the interaction between V_{SD} and the reference function, $V_{pqrs} = \langle\Psi_0|\hat{H}|pqrs\rangle$. This equation cannot be solved directly due to the linear dependence in the basis. Once the linear dependences are removed, the equation has a somewhat smaller dimension and is diagonalized using some iterative approach, usually the Davidson diagonalization method.

The Fock matrix is diagonalized in blocks corresponding to inactive/inactive, active/active, and external/external submatrices. Unlike in the CASPT2 model, in RASPT2 the active/active submatrix is divided itself into three subblocks defined by RAS1, RAS2, and RAS3 orbital spaces. Orbitals rotations that couple different RAS subspaces are not allowed and diagonalization of the active part of the Fock matrix (3×3 block structure) is not longer complete in RASPT2 as it is in CASPT2 (one block structure).

In several cases, some corrections have to be applied over the zeroth order Hamiltonian H_0 because of the presence of intruder states when excited states calculations are carried out [10, 11, 12] and owing to some unbalanced description of H_0 for closed-shell electronic structures against open-shell ones [10].

Intruder states are states in the V_{SD} space that have a zeroth order energy which is close or even below the reference energy E_0. In Møller-Plesset perturbation theory to Second Order (MP2) such states rarely occur, since there is a large separation in energy between occupied and external molecular orbitals. However, in the CASPT2 method, this is not always the case. The diagonal elements of the Fock matrix are related to ionization energies (IP) for inactive orbitals and to electron affinities (EA) for external orbitals. For active orbitals we have orbital energies resulting from an average between the IP and the EA. Weakly occupied orbitals may thus have energies close to those of external orbitals and strongly occupied orbitals close to inactive orbitals. Since active orbitals may be both excited out from and into, it may happen that the energy denominators become small. If this happens, there is an intruder state in the V_{SD} space. The corresponding first order coefficient may then become large and a treatment based on perturbation theory is no longer valid.

Two cases of intruder states can be found, strong and weak intruder states. The strong intruder states are characterized by a large interaction with the reference function and then the contribution of the second order energy cannot be neglected. The only solution is to increase the active space to include the orbitals which cause the low denominators. If this cannot be done because of the size limit on the active space, second order perturbation theory cannot be used to treat the problem. Weak intruder states are the most common case. Their interaction with the reference state is small, so the contribution to the correlation energy can be expected to be small. It is then possible to remove the intruder state by a level shift (LS) technique [12]. A constant, ε, is added to the zeroth order Hamiltonian yielding a modified first order equation:

$$\left(\hat{H}_0 - E_0 + \varepsilon\right)|\tilde{\Psi}_1\rangle = \left(E_1 - \hat{H}_1\right)|\Psi_0\rangle \tag{1.12}$$

where the tilde indicates that a level shifted zeroth order Hamiltonian has been used. If it is assumed that the zeroth order Hamiltonian is diagonal in

the expansion space, Φ_μ, the second order energy for the shifted Hamiltonian, $\tilde{E}_2$, can be expressed as:

$$\tilde{E}_2 = E_2 + \varepsilon \sum_\mu |\tilde{c}_\mu|^2 \left(1 + \frac{\varepsilon}{\epsilon_\mu - E_0}\right) \tag{1.13}$$

where E_2 is the second order energy obtained without a level shift and ϵ_μ is the sum of the orbital energies of the spin-orbitals occupied in the given configuration. Assuming that $\epsilon_\mu - E_0 >> \varepsilon$ and neglecting terms quadratic in ε, it is obtained:

$$E_2 = \tilde{E}_2 - \varepsilon \left(\frac{1}{\tilde{\omega} - 1}\right) \tag{1.14}$$

where unshifted second order energy is expressed in terms of the quantities obtained with the level shift, $\tilde{E}_2$ and $\tilde{\omega}$ (weight of the reference wave function). In the absence of an intruder state the back shifted second order energy in 1.14 depends only very weakly on the value of the level shift.

In the standard Fock matrix formulation of the zeroth order Hamiltonian, bond energies are normally underestimated with between 2 and 5 kcal·mol^{-1} for each bond formed. This is due to a systematic error in the method [13]. In every process where the number of paired electrons is changed, an error of this size will occur for each electron pair. The same error affects excitation energies, which are often a little bit too small. The errors are due to an unbalanced description of the zeroth order Hamiltonian for a closed shell electronic structure in comparison to structures with unpaired electrons.

In the closed shell case, the diagonal elements of $\mathbf{F}$ will correspond to $-IP$ (IP, ionization energy) when the orbital is innactive and to $-EA$ (EA, electron affinity) when it is virtual. For active orbitals, the diagonal elements of $\mathbf{F}$ will

be a weighted average of $-IP$ and $-EA$. Let us attempt to write Eq. 1.8 as an interpolation of the two extreme cases:

$$f_{pp} = -\frac{1}{2}\left(D_{pp}(IP)_p + (2 - D_{pp})(EA)_p\right) \tag{1.15}$$

where D_{pp} is the diagonal element of the one-particle density matrix for orbital p. This formula is correct for $D_{pp} = 0$ and 2, and also for a singly occupied open shell. Thus, for an open shell ($D_{pp} = 1$) we obtain:

$$f_{pp} = -\frac{1}{2}\left((IP)_p + (EA)_p\right) \tag{1.16}$$

This feature of the Fock operator will lead to denominators in the expression for the second order energy that are too small in the case of excitation into or out of a partially occupied orbital. The energy of systems with open shells will be too low, so the perturbation energy is overestimated. A second problem arising from this formulation is an increased possibility for intruder states.

These deficiencies in the zeroth order Hamiltonian can be solved by adding a shift to Eq. 1.15. Let us assume that when we excite into an active orbital, we want its energy to be replaced by $-EA$. In this case, the shift is:

$$\sigma_p^{(EA)} = \frac{1}{2}D_{pp}\left((IP)_p - (EA)_p\right) \tag{1.17}$$

On the other hand, if we excite out of this orbital we want its energy to be replaced by $-IP$ and the corresponding shift is:

$$\sigma_p^{(IP)} = -\frac{1}{2}(2 - D_{pp})\left((IP)_p - (EA)_p\right) \tag{1.18}$$

Unfortunately, the definition of $(IP)_p$ and $(EA)_p$ is not straightforward. Therefore, we replace $(IP)_p - (EA)_p$ with an average shift parameter ϵ. The two shifts are then:

$$\sigma_p^{(EA)} = \frac{1}{2} D_{pp} \epsilon \tag{1.19}$$

$$\sigma_p^{(IP)} = -\frac{1}{2}(2 - D_{pp})\epsilon \tag{1.20}$$

where ϵ would have to be considered as a shift parameter, to be determined by comparison to accurate calculations or eventually experimental results. The overestimation of the correlation energy in open shell cases is corrected, for instance, introducing a shifted IPEA parameter [14] which takes as optimal value 0.25 au.

These two last tools, LS and IPEA, are available both in the CASPT2 and RASPT2 models.

A final problem has to be mentioned in the RASPT2 formulation. By RAS constraints, only RAS1 and RAS3 subspaces should handle dynamical correlation whereas RAS2 should take into account all non-dynamical correlation effects. Unfortunately, this subdivision is not well defined. Therefore, to address all dynamical correlation effects in the RASPT2, a perturbative treatment of the fully internal excitations must be included. This implies in the present formulation of H_0 building fourth order density matrices and this is unpractical with larger active spaces. New approximations on H_0 are needed to solve this question, which, in any case its importance has to be evaluated.

As it occurs with single-state CASPT2 calculations, in RASPT2 the states computed individually from a previous CASSCF reference are not orthogonal to the other states of the same spatial and spin symmetry. As a consequence the obtained solutions have ignored the coupling between the states. The Multi-state CASPT2 (MS-CASPT2) [15] procedure represents an extension

of the CASPT2 method for the perturbation treatment of chemical situations that require two or more reference states. For instance, situations such as avoided crossings and near-degeneracy of valence and Rydberg states, and their proper description cannot be fully accounted for by just using a single-state perturbation treatment.

In the MS-CASPT2 or MS-RASPT2 methods an effective Hamiltonian matrix is constructed where the diagonal elements correspond to the CASPT2 energies and the off-diagonal elements introduce the coupling to second order in the dynamical correlation energy. Let us assume that we have performed N CASPT2 calculations for the corresponding reference wave functions Ψ_{0i} $(i = 1, N)$, obtained by using average CASSCF for those N roots and a set of average molecular orbitals is, therefore, available. In order to build the matrix representation of the Hamiltonian using as basis set the N normalized wave functions corrected up to first order, $|\Psi_i\rangle = |\Psi_{0i}\rangle + |\Psi_{1i}\rangle$, the following matrices are defined:

$$S_{ij} = \langle\Psi_i|\Psi_j\rangle = \langle\Psi_{0i} + \Psi_{1i}|\Psi_{0j} + \Psi_{1j}\rangle = \delta_{ij} + s_{ij} \tag{1.21}$$

$$\langle\Psi_{0i}|\hat{H}|\Psi_{0j}\rangle = \delta_{ij}E_i \tag{1.22}$$

$$\langle\Psi_{0i}|\hat{H}|\Psi_{1j}\rangle = e_{ij} \tag{1.23}$$

Notice that the two wave functions are not orthogonal, since $\langle\Psi_{0i}|\Psi_{0j}\rangle = \delta_{ij}$ and $\langle\Psi_{0i}|\Psi_{1j}\rangle = 0$ but $\langle\Psi_{1i}|\Psi_{1j}\rangle = s_{ij}$. On the other hand, the CASSCF energy for state i is represented by E_i and the elements e_{ij} are the CASPT2 correlation energies. For each state, the Hamiltonian can be expressed as the sum of a zeroth order contribution and a Hamiltonian taking care of the remaining effects, $\hat{H} = \hat{H}_i^0 + \hat{H}_i'$. Therefore, up to second order it holds true that:

$$\langle\Psi_{1i}|\hat{H}|\Psi_{1j}\rangle \approx \langle\Psi_{1i}|\hat{H}_i^0|\Psi_{1j}\rangle \approx \langle\Psi_{1i}|\hat{H}_j^0|\Psi_{1j}\rangle \tag{1.24}$$

The elements $\langle\Psi_{1i}|\hat{H}_i^{'}|\Psi_{1j}\rangle$ correspond to a third order correction and, consequently, they are not considered. The matrix representation of the Hamiltonian is not symmetric. Assuming that the off-diagonal terms are very similar, as it is implicit from Eq. 1.24, the matrix is made symmetric by using the average value:

$$\langle\Psi_{1i}|\hat{H}|\Psi_{1j}\rangle = \frac{1}{2}\left(\langle\Psi_{1i}|\hat{H}_i^0|\Psi_{1j}\rangle + \langle\Psi_{1i}|\hat{H}_j^0|\Psi_{1j}\rangle\right) \tag{1.25}$$

The matrix element including zeroth, first, and second order corrections takes the general form:

$$\hat{H}_{ij} = \langle\Psi_i|\hat{H}|\Psi_j\rangle = \delta_{ij}E_i + \frac{1}{2}(e_{ij} + e_{ji}) + \frac{1}{2}\left(E_i^{(0)} + E_j^{(0)}\right)s_{ij} \tag{1.26}$$

By solving the corresponding secular equation $(\mathbf{H} - E\mathbf{S})\,\mathbf{C} = 0$, the eigenfunctions and eigenvalues can be obtained. They correspond to the MS-CASPT2 wave functions and energies, respectively.

The MS-CASPT2 wave function can be finally written as:

$$|\Psi_p\rangle = \sum_i c_{pi}|\Psi_{0i}\rangle + |\Psi_{1p}\rangle = |\Psi_{0p}\rangle + |\Psi_{1p}\rangle \tag{1.27}$$

where $|\Psi_{0i}\rangle$ are the CASSCF reference functions and $|\Psi_{1p}\rangle$ is the first order wave function for state p. Accordingly, the function $|\Psi_{0p}\rangle$ formed by a linear combination of the CAS states involved in the MS-CASPT2 calculation is the model state and can be considered as a new reference function for state p. The reference functions $|\Psi_{0p}\rangle$ are the so-called Perturbation Modified CAS (PMCAS) [15]. They are used for the computation of transition properties and expectation values at the MS-CASPT2 level.

For the proper use of the MS-CASPT2 method, the condition of Eq. 1.24 has to be fulfilled. In practice, it means that the asymmetric effective Hamiltonian matrix should have small and similar off-diagonal elements. Otherwise, the average process carried out, $(H_{12} + H_{21})/2$, may lead to unphysical results in both the MS-CASPT2 energies and eigenfunctions. The condition that $H_{12} \approx H_{21}$ can be achieved by enlarging the active space, which implies a redefinition of the zeroth order Hamiltonian. Large active spaces, beyond the main valence MOs, are used naturally in the simultaneous treatment of valence and Rydberg states, where the MS-CASPT2 approach has proved to be extremely useful. Especial caution has to be exercised, however, for the computation of a crossing point between two surfaces, as in the case of conical intersections (and avoided crossings), crucial in photochemistry [16].

1.2. The Hubbard model

In the fifties much attention was given to the theory of correlation effects in the free electron gas which is used as a model for the conduction bands of metals and alloys [17, 18, 19]. Nevertheless, transition and rare-earth metals have, apart from their conduction band, partially filled d- or f- bands which require correlation phenomena to determine the properties of these narrow energy bands. The free electron gas does not provide a good model for these bands. Rather, one requires a theory of correlations which takes into account adequately the atomistic nature of the solid. John Hubbard (1963) proposed a model to determine the balance between band-like and atomic-like behavior in d-bands of transition metals [20]. Basically, the Hubbard model features electrons that can hop between *sites*, representing ions, that are arranged in an ordered, crystalline pattern. When two electrons are on the same site, they have to pay an energy penalty due to their mutual repulsion. This introduces additional correlations between the electrons beyond those due to the Pauli exclusion principle. If the electron-electron repulsion is weak, then the electrons roam more or less freely around the material and it behaves like a metal. But if it is strong, they are forced to localize at fixed atomic positions and the material behaves like an insulator. In a *half filled* system (one electron per site), the model predicts the formation of a Mott insulator like some transition metal oxides. When the Mott insulator is doped, by either adding or removing a small number of electrons, it becomes metallic [21].

A prototypical metal in condensed matter physics might be described by the Hamiltonian [22]:

$$
\begin{aligned}
H &= H_e + H_i + H_{ei} \\
H_e &= \sum_i \frac{\mathbf{p}_i^2}{2m} + \sum_{i \neq j} V_{ee}(\mathbf{r}_i - \mathbf{r}_j) \\
H_i &= \sum_I \frac{\mathbf{P}_I^2}{2m} + \sum_{I \neq J} V_{ii}(\mathbf{R}_I - \mathbf{R}_J) \\
H_{ei} &= \sum_{iI} V_{ei}(\mathbf{R}_I - \mathbf{r}_i)
\end{aligned}
\tag{1.28}
$$

where $\mathbf{r}_i(\mathbf{R}_I)$ are the coordinates of $i, i = 1, 2, ..., N$, valence electrons (ion cores) and H_e, H_i and H_{ei} describe the dynamics of the electrons, ions and the interaction of electrons and ions, respectively.

In order to study the dynamics of the conduction electrons, the full Hamiltonian is reduced to the pure electronic Hamiltonian H_e and the interaction between electrons and the positively charged ionic background lattice H_{ei}. Using the second quantization notation, the Hamiltonian used to describe the electronic system is:

$$
\begin{aligned}
H = H_0 + V_{ee} = \int \mathrm{d}\mathbf{r}\, a_\sigma^\dagger(\mathbf{r}) \left[\frac{\mathbf{p}^2}{2m} + V(\mathbf{r}) \right] a_\sigma(\mathbf{r}) + \\
+ \frac{1}{2} \int \mathrm{d}\mathbf{r} \int \mathrm{d}\mathbf{r}'\, a_\sigma^\dagger(\mathbf{r}) a_\sigma(\mathbf{r}) V(\mathbf{r} - \mathbf{r}') a_\sigma^\dagger(\mathbf{r}') a_\sigma(\mathbf{r}')
\end{aligned}
\tag{1.29}
$$

where $V(\mathbf{r}) = \sum_I V_{ei}(\mathbf{R}_I - \mathbf{r})$ and the coordinates of the lattice ions are presumed fixed. For completeness, the creation ($a_\sigma^\dagger(\mathbf{r})$) and annihilation ($a_\sigma(\mathbf{r})$) operators are labeled with a spin index $\sigma = \pm 1/2$. The Hamiltonian 1.29 defines the problem of the interacting electron gas embedded into a solid state system.

Considering the single-particle Hamiltonian of electrons in a periodic potential and vanishing V_{ee} interaction, the eigenstates of the Hamiltonian according to the Bloch's theorem are represented in the form of Bloch waves:

$$\psi_{\mathbf{k}n}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}n}(\mathbf{r}) \tag{1.30}$$

where $\mathbf{k}$ are the components of the crystal momentum, $k_i \in [-\pi/a, \pi/a]$ being a the lattice parameter, and it is assumed the periodicity of the lattice potential and the Bloch functions in all directions, $V(\mathbf{r}+a\mathbf{e}_i) = V(\mathbf{r})$ and $u_{\mathbf{k}n}(\mathbf{r}+a\mathbf{e}_i) = u_{\mathbf{k}n}(\mathbf{r})$, and n labels the energy bands.

There are two classes of materials where the Bloch functions are modified when V_{ee} is added to the Hamiltonian: nearly free electron systems and tight binding systems. The interest of the present thesis is focused on the second group of materials.

Assuming that the electrons are tightly bound to the lattice centres, it is convenient to expand the wave function describing the system in terms of the local basis of **Wannier states**:

$$|\psi_{\mathbf{R}n}\rangle \equiv \frac{1}{\sqrt{N}} \sum_{\mathbf{k}}^{B.Z.} e^{-i\mathbf{k}\cdot\mathbf{R}} |\psi_{\mathbf{k}n}\rangle, \qquad |\psi_{\mathbf{k}n}\rangle = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} |\psi_{\mathbf{R}n}\rangle \tag{1.31}$$

where $\mathbf{R}$ are the coordinates of the lattice centres and the summation goes over all momenta $k_i = n_i \frac{2\pi}{L} \in [0, 2\pi/a]$ in the first Brillouin zone. For a system with vanishing weak interatomic overlap, i.e. a lattice where V approaches a superposition of independent atomic potentials, the Wannier function $\psi_{\mathbf{R}n}(\mathbf{r}) \equiv \langle \mathbf{r}|\psi_{\mathbf{R}n}\rangle$ equals the nth orbital of an isolated atom centred at coordinate $\mathbf{R}$. However, for finite interatomic coupling, the N degenerate states labeled by n split to form an energy band (see Fig. 1.2). In this

case, the Wannier functions differ from isolated atomic orbitals. Unlike the Bloch functions, the Wannier functions are not eigenfunctions of the electronic Hamiltonian.

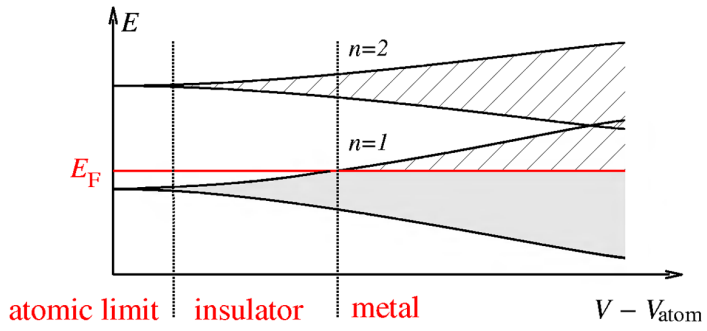


Figure 1.2: Energy bands in a solid. For a weakly bound solid, the energy levels are N-fold degenerate and coincide with the bound state levels of isolated atom-potentials V_{atom}. For finite interatomic coupling, levels split to form bands.

In cases where the Fermi energy E_{F} is situated between two energetically separated bands, the material is an insulator. If the Fermi energy is, however, located within a band, the system is metallic. Ignoring the complications that arise when bands begin to overlap, henceforth, systems where the Fermi energy is found within a definite band n_0 are considered.

In order to express the electronic Hamiltonian 1.29 in the Wannier-language, it must be noticed that the Wannier states $\{|\psi_{\mathbf{R}}\rangle\}$ form an orthonormal basis of the single particle Hilbert space, i.e. the transformation between the real-space and the Wannier representation:

$$|\mathbf{r}\rangle = \sum_{\mathbf{R}} \langle \psi_{\mathbf{R}} | \mathbf{r} \rangle | \psi_{\mathbf{R}} \rangle = \sum_{\mathbf{R}} \psi^*_{\mathbf{R}}(\mathbf{r}) | \psi_{\mathbf{R}} \rangle \tag{1.32}$$

is unitary (the n band-index has been dropped because only contribution coming from the n_0 band are of interest). As such, it induces a transformation:

$$a^\dagger_\sigma(\mathbf{r}) = \sum_{\mathbf{R}} \psi^*_{\mathbf{R}}(\mathbf{r}) a^\dagger_{\mathbf{R}\sigma} \equiv \sum_i \psi^*_{\mathbf{R_i}}(\mathbf{r}) a^\dagger_{i\sigma} \tag{1.33}$$

between the real space and the Wannier state operator basis. In the last term of the expression 1.33, the lattice centre coordinates are labeled as $\mathbf{R} \equiv \mathbf{R}_i$ by counting index $i = 1, 2, ..., N$. Similarly, the unitary transformation between Bloch and Wannier states, 1.31, induces an operator transformation:

$$a^\dagger_{\mathbf{k}\sigma} = \frac{1}{\sqrt{N}} \sum_i e^{i\mathbf{k}\cdot\mathbf{R}_i} a^\dagger_{i\sigma} \tag{1.34}$$

Using both formulae 1.33 and 1.34 to write the Hamiltonian 1.29 in the Wannier formulation and taking into account that Bloch states diagonalize the single particle Hamiltonian H_0, it is obtained:

$$\begin{aligned} H_0 = \int \mathrm{d}\mathbf{r}\, a^\dagger_\sigma(\mathbf{r}) \left[\frac{\mathbf{p}^2}{2m} + V(\mathbf{r})\right] a_\sigma(\mathbf{r}) &= \sum_{\mathbf{k}} a^\dagger_{\mathbf{k}\sigma} \epsilon_k a_{\mathbf{k}\sigma} \\ &= \frac{1}{N} \sum_{ii'} \sum_{\mathbf{k}} e^{i\mathbf{k}(\mathbf{R}_i - \mathbf{R}_{i'})} \epsilon_k a^\dagger_{i\sigma} a_{i'\sigma} \equiv \sum_{ii'} a^\dagger_{i\sigma} t_{ii'} a_{i'\sigma} \end{aligned} \tag{1.35}$$

where the σ index is dropped in the summations and the term $t_{ii'}$ is defined as $t_{ii'} = \frac{1}{N}\sum_k e^{i\mathbf{k}(\mathbf{R}_i - \mathbf{R}_{i'})}\epsilon_k$. The new representation of H_0 describes **electron hopping** from one lattice centre i to another. The strength of the hopping amplitude $t_{ii'}$ is controlled by the effective overlap of neighboring atoms. In the atomic limit, where the levels $\epsilon_k =$ constant are degenerate, $t_{ii'} = \delta_{ii'}$ and

no interatomic transport is possible. The tight binding representation is useful if $t_{i\neq i'}$ is non-vanishing but so weak that only the nearest neighbor hopping effectively contributes.

Considering V_{ee} and using the expression 1.33, it is obtained:

$$V_{ee} = \sum_{ii'jj'} U_{ii'jj'}\, a^\dagger_{i\sigma} a_{j\sigma} a^\dagger_{i'\sigma'} a_{j'\sigma'} \tag{1.36}$$

where the matrix element $U_{ii'jj'}$ is defined as:

$$U_{ii'jj'} = \frac{1}{2} \int \mathrm{d}\mathbf{r}\, \mathrm{d}\mathbf{r}'\, \psi^*_{\mathbf{R}_i}(\mathbf{r}) \psi_{\mathbf{R}_j}(\mathbf{r}) V(\mathbf{r}-\mathbf{r}') \psi^*_{\mathbf{R}_{i'}}(\mathbf{r}') \psi_{\mathbf{R}_{j'}}(\mathbf{r}') \tag{1.37}$$

The combination of the contributions H_0 and V_{ee} is shown below:

$$H = \sum_{ii'} a^\dagger_{i\sigma} t_{ii'} a_{i'\sigma} + \sum_{ii'jj'} U_{ii'jj'}\, a^\dagger_{i\sigma} a_{j\sigma} a^\dagger_{i'\sigma'} a_{j'\sigma'} \tag{1.38}$$

and defines the **tight binding representation** of the interaction Hamiltonian.

Assuming electrons are tightly bound, the contributions of $U_{ii'jj'}$ can be identified:

- **Direct terms** $U_{iii'i'} \equiv V_{ii'}$ accounts for the essentially classical interaction between charges localized at neighboring sites.

$$\sum_{i\neq i'} V_{ii'} \hat{n}_i \hat{n}_{i'} \tag{1.39}$$

where $\hat{n}_i = \sum_\sigma a^\dagger_{i\sigma} a_{i\sigma}$ is the occupation number operator.

- **Exchange coupling** which induce magnetic correlations among the electronic spins. Setting $J_{ij} \equiv U_{ijji}$ and making use of the Pauli matrix identity $\sigma_{\alpha\beta} \cdot \sigma_{\gamma\delta} = 2\delta_{\alpha\delta}\delta_{\beta\gamma} - \delta_{\alpha\beta}\delta_{\gamma\delta}$ and the relation for the spin operator $\hat{\mathbf{S}}_m = a^\dagger_{m\alpha}\sigma_{\alpha\beta}a_{m\beta}/2$, the following expression is obtained:

$$\sum_{i\neq j} U_{ijji}\, a^\dagger_{i\sigma} a_{j\sigma} a^\dagger_{j\sigma'} a_{i\sigma'} = -2 \sum_{i\neq j} J_{ij} \left(\hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j + \frac{1}{4}\hat{n}_i\hat{n}_j \right) \tag{1.40}$$

  Equation 1.40 resembles the Heisenberg spin Hamiltonian. Such contributions tend to induce weak **ferromagnetic coupling** of neighboring spins. Consider two electrons inhabiting neighboring sites, the Coulomb repulsion between the particles is minimized if the orbital-two-particle wave function is antisymmetric and, therefore, has low amplitude in the interaction zone between the particles. Since the overall wave function must be antisymmetric, the spin wave function is enforced to be symmetric, i.e. the two spins must be aligned.

- **Atomic limit** where the atoms are well-separated and the overlap between neighboring orbitals is weak, the matrix elements t_{ij} and J_{ij} are exponentially small in the interatomic separation. In the limit, the 'on-site' Coulomb or Hubbard interaction, U_{iiii}, is expressed as:

$$\sum_{i\sigma} U_{iiii}\, a^\dagger_{i\sigma} a_{i\sigma} a^\dagger_{i\bar{\sigma}} a_{i\bar{\sigma}} = \sum_{i\sigma} \left(U \hat{n}_{i\sigma}\hat{n}_{i\bar{\sigma}} - \frac{U}{2}\hat{n}_i \right) \tag{1.41}$$

  where $U \equiv 2U_{iiii}$ generates the dominant interaction mechanism.

The Hubbard Hamiltonian is defined as:

$$H = \sum_{ii'\sigma} a^{\dagger}_{i\sigma} t_{ii'} a_{i'\sigma} + U \sum_{i\sigma} \hat{n}_{i\sigma} \hat{n}_{i\bar{\sigma}} \tag{1.42}$$

where the term $\frac{U}{2} \sum_i \hat{n}_i$ has been omitted so it generates a trivial shift of the chemical potential. The summation extends over the nearest neighbor lattice sites and the spin σ index is also indicated. The second term in Eq. 1.42 provides an excess repulsion U, relative to a configuration with one electron per site, whenever two electrons occupy the same site.

The Hubbard Hamiltonian is not only used in solid state physics but also in chemistry as a model of π-conjugated hydrocarbons [23, 24]. The Hubbard model is reduced exactly to the Hückel Molecular Orbital (HMO) model when U is equal to zero, while as $lim\, U/t \to \infty$, it is equivalent to the Heisenberg spin Hamiltonian or the nearest-neighbor valence bond (VB) model.

The Hückel model represents one of the simplest attemps to describe the most important effects of chemical bonding. The HMO retains two types of terms: a site energy α_i which represents the attraction of an electron in atomic orbital i to its atom, and a bonding parameter $\beta_{i,i'}$ which measures the strength of a covalent bond between atomic orbital i and atomic orbital i'. These both quantities corresponds to t_{ii} and $t_{ii'}$, respectively, in the Hubbard model. Nevertheless, the Hückel model suffers from two serious limitations. First, it predicts the singlet and triplet transition to excited states that share the same initial and final molecular orbitals to be degenerate. On the other hand, there is no explicit electron-electron interaction in the Hamiltonian thus neglecting electron correlation. The Hubbard model represents an extension of the Hückel model and goes continuously from the molecular orbital to the valence bond limit as the strength of the electron-electron interaction is varied.

Rather than limiting electron-electron interactions to one centre, the extended Hubbard model adds an electron repulsion term V for two electrons on nearest-neighbor sites. The equation 1.42 can be rewritten as

$$H = \sum_{ii'\sigma} a^{\dagger}_{i\sigma} t_{ii'} a_{i'\sigma} + U \sum_{i\sigma} \hat{n}_{i\sigma} \hat{n}_{i\bar{\sigma}} + \frac{1}{2} \sum_{ii'} V_{ii'} \left(Z_i - \hat{n}_i\right) \left(Z_{i'} - \hat{n}_{i'}\right) + V_c \quad (1.43)$$

where the double sums run over nearest-neighbors, Z_i represents the core charges of the atoms and V_c accounts for the effects of the nuclei and core electrons not considered explicitly in the model. The extended Hubbard model is analytically solvable for a system consisting of just two identical sites and two electrons, giving the energy eigenvalues for the three lowest energy states - ground state (G), and the lowest excited triplet (T) and singlet (S) states - as follows:

$$\begin{aligned} E_G(R) &= -\frac{1}{2} \left\{ \left[(U-V)^2 + 16t^2 \right]^{1/2} - (U-V) \right\} + V_c(R) \\ E_T(R) &= V_c(R) \\ E_S(R) &= (U-V) + V_c(R) \end{aligned} \quad (1.44)$$

For systems in which all of the atoms are chemically identical, the Hamiltonian term containing t_{ii} adds a constant to the energy of all states at all geometries and it may be set to zero. The indices for $t_{ii'}$ and $V_{ii'}$ have been dropped since there is just one bond and one type of site. The term V depends on the distance R between the nuclei, while U can be taken as essentially an atomic property. For simplicity, $U - V$ is referred as U^{eff}, an effective repulsion potential between the two active electrons.

A two-site two-electron π-conjugated system is the ethene molecule where each carbon atom is presumed to form three localized sigma bonds, whose energy is incorporated into V_c, and to possess one active $2p_z$-type atomic orbital perpendicular to the molecular plane which is active in the Hubbard model.

All parameters of the extended Hubbard model can be extracted from the potential energy curves of the two-site system since U is given by the large-R limit of U^{eff}, where V falls to zero, and $V(R)$ can then be found from the distance dependence of U^{eff}. Nevertheless, the simple three-state picture of the low energy spectrum breaks down long before V becomes small because the excited singlet dissociates to an excited state of the separated fragments. Thus, for applications to conjugated hydrocarbons, U has been fixed in the extended model at 5.0 eV and then V is extracted as a function of R from the $E_S(R) - E_T(R)$ excited state potential curves. Fortunately, the performance of the model depends only weakly on the exact value chosen for U.

1.3. Boron Hydrides

In 1916, G. N. Lewis was the first to interpret the Couper lines between the chemical symbols (1858) as bonds of electron pairs [25]. Since the publication of the Schrödinger equation in 1926 [26], the theories of bond first started with the work of Heitler and London on the H_2 molecule (1927) [27], and Pauling extended the theory with the concepts of resonance and orbital hybridization to derive the valence bond (VB) theory [28]. The stoichiometry postulated by van't Hoff and Le Bel for organic compounds, and by Werner for inorganic complexes, found their quantum-mechanical explanation in the hybridization theory (Pauling, 1931). Together with Nyholm and Gillespie's VSEPR (Valence Shell Electron Pair Repulsion) theory (1957) [29], the above considerations were in accordance with the X-ray structures of thousands of ionic and molecular species. By use of the valence bond theory, Pauling related the bond lengths to the bond order and the electronegativity to the bond polarization.

During the same period as G. N. Lewis proposed the electron-pair bonds, in a remarkable series of papers from 1912 to 1936, Alfred Stock and co-workers prepared a set of boron compounds known as boranes or boron hydrides which were characterized as very reactive against O_2. Stock and co-workers designed experiments *in vacuo* to manipulate volatile boron hydrides without O_2 and humidity. Stock characterized the following boranes: B_2H_6, B_4H_{10}, B_5H_9, B_5H_{11}, B_6H_{10} and $B_{10}H_{14}$. In Fig. 1.3 the structures of boron hydrides discovered by Stock are shown [30]. Each boron atom is bonded by at least one hydrogen atom and, sometimes, hydrogen atoms establish hydrogen bridge bonds between two boron atoms. The reactivity of the boranes is extremely diverse and some chemical reactions are: (a) acidity and the formation of borane anions, (b) substitution of H by halogens, organic groups, among others, (c) adduct formation, (d) electron addition to give anions, and (e) skeletal additions and formation of compounds that have heteroatoms (C, Si, P, etc.) in the skeleton [31].

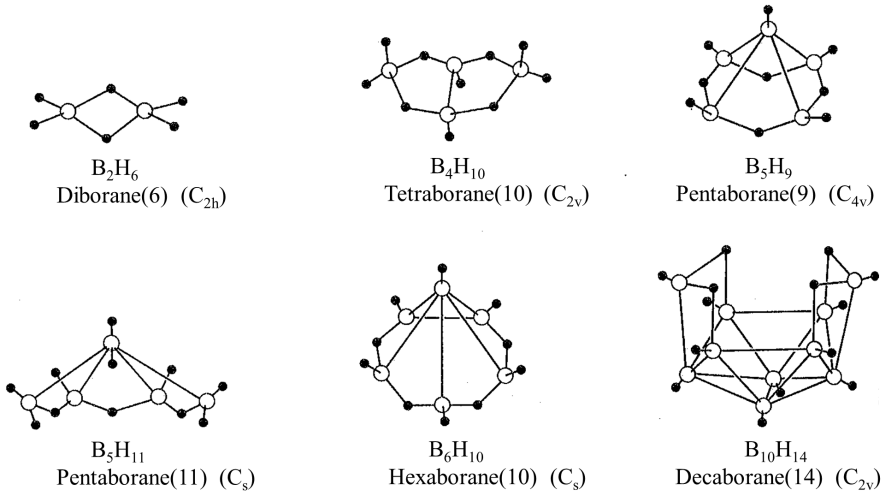


Figure 1.3: Stock's structures of boron hydrides.

Regarding structural and bonding features, boron clusters are molecular systems composed of boron atoms placed in vertices of triangular-faced polyhedra (deltahedra). Each boron atom can contribute with a maximum number of three electrons to their nearest neighbours, which can be up to five, six or seven atoms. Therefore, the connexions between atoms cannot be considered as single bonds of an electron pair. This problem is known as *electron deficiency* and it is managed by the use of the *multicenter bonding* concept. Several structural-bonding elements are used to describe the different topological structures of boranes: two-center two-electron BH groups, three-center two-electron BHB groups, two-center two-electron BB groups, three-center two-electron open BBB groups, and three-center two-electron closed BBB groups. For the boranes themselves the open three-center two-electron boron bridge bond is not normally required, but it is used in some heterosystems [30, 31].

Single-cluster boranes are either most spherical deltahedra, i.e. *closo*-boranes, or are derived by the removal of one, two, or three vertices of high connectivity from *closo*-structures to give *nido*-, *arachno*- and *hypho*-, respectively [32]. In

Fig. 1.4 a scheme of these cluster shapes connected between them by diagonal arrows is shown [30, 31, 33]. The stoichiometry of the different structural clusters are in the head of the table. From the *closo*-borane dianions one can obtain neutral open-faced boranes depending on the number of the removed vertices. There is a relation between the geometry of boron hydrides and their electronic structure established by the Wade's rule [34]. For a borane with p vertices, the number of skeletal electron pairs depends on the shape of the cluster as shown in Table 1.1.

Table 1.1: Relation between single-cluster types and their skeletal electron pairs.

Type of the structure	**Skeletal electron pairs**
closo	$p+1$
nido	$p+2$
arachno	$p+3$
hypho	$p+4$

Large structures are composed of smaller clusters with two different modes of combining individual single clusters: (i) a two-center two-electron or three-center two-electron bond connects two independent units, or (ii) one, two, three or four vertices are shared by two individual units. The resulting clusters of fused polyhedral units have been named macropolyhedra [32].

Jemmis' *mno* rule, a skeletal electron count principle [35], can be easily and correctly employed to any macropolyhedral borane. In short, the sum of the number of single-cluster fragments (m), the number of vertices in the macropolyhedron (n), the number of single-vertex-sharing junctions (o), and the number of missing vertices (p) equals the number of skeletal electron pairs of a macropolyhedral borane.

A large number of known macropolyhedral boranes with the general formula B_nH_{n+4} consist of two *nido* units sharing two vertices. Formally, two vertex-sharing *nido:nido*-macropolyhedra derive from the condensation reaction of

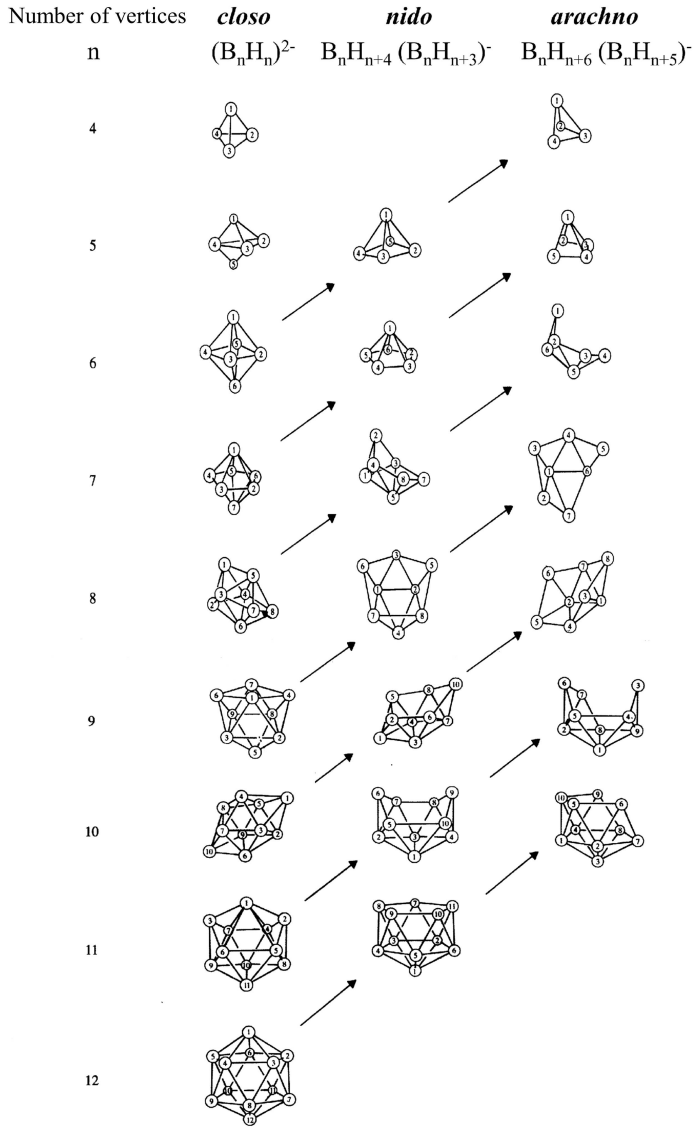
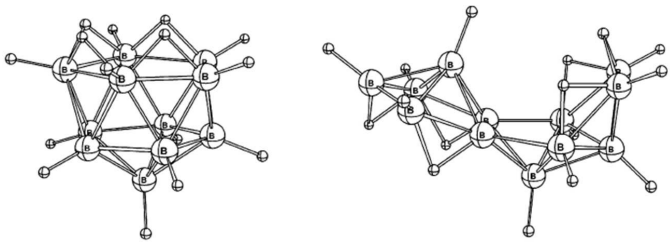


Figure 1.4: Polyhedral structures of single-cluster boranes with diagonal arrows indicating the *closo*-, *nido*- and *arachno*- structures connected by successive removal of BH vertex of high connectivity. Hydrogens are not shown.

two *nido*-boranes releasing B_2H_6. B_nH_{n+4} macropolyhedra with one or three vertices shared between two *nido* fragments do not obey the *mno* rule and are experimentally unknown. The largest experimentally known homonuclear *nido* single cluster is $B_{11}H_{15}$ [36], whereas the smallest experimentally known homonuclear *nido:nido*-macropolyhedral borane is $B_{12}H_{16}$ [37, 38] (Fig. 1.5). The latter consists of one eight-vertex *nido* unit sharing two vertices with another six-vertex *nido* fragment [32].



nido-$B_{11}H_{15}$ *nido*(6):*nido*(8)-$B_{12}H_{16}$

Figure 1.5: The largest experimentally known homonuclear *nido* single cluster, $B_{11}H_{15}$, and the smallest experimentally known homonuclear *nido:nido*-macropolyhedral borane cluster, $B_{12}H_{16}$.

Any neutral *nido:nido*-macropolyhedral borane is less stable than the corresponding *nido* single-cluster borane as long as the total number of vertices is equal to or less than 11. For 12 or more vertices, up to 19, macropolyhedra exist that are energetically preferred over the single-cluster alternatives. The increase in the stability of *nido:nido*-macropolyhedral boranes can be explained on the basis of the connectivity of boron vertices. Five-coordinate vertices are especially favorable, as indicated by the high stability of icosahedral *closo*-$[B_{12}H_{12}]^{2-}$. Large-sized *nido* single-cluster boranes usually have more highly connected vertices in addition. Isomeric *nido:nido*-macropolyhedral boranes, on the other hand, are built from smaller cluster fragments and have a smaller number of highly connected vertices (usually more five-coordinate vertices), and therefore get enhanced stability [32].

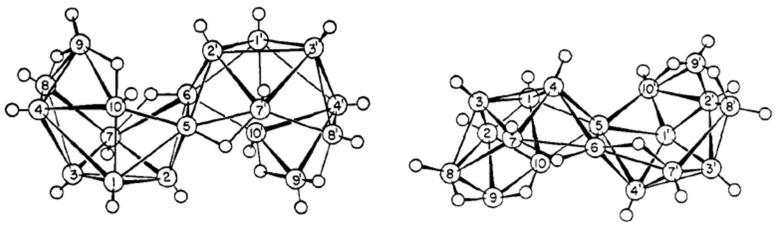
Hydrogen bonds at the open face prefer to bridge vertices of least connectivity. Vertices at the open face of *nido:nido*-macropolyhedral boranes are usually less connected as compared to vertices at the open face of corresponding single *nido* clusters.

nido(10):*nido*(10)-octadecaborane(22) $B_{18}H_{22}$ isomers

In the early sixties, two isomeric $B_{18}H_{22}$ hydrides were prepared [39] and characterized by X-ray diffraction studies [40, 41, 42, 43]. The two isomers of $B_{18}H_{22}$ (*anti* and *syn*) were formed simultaneously during the hydrolisis of the hydronium ion salt of $[B_{20}H_{18}]^{2-}$.

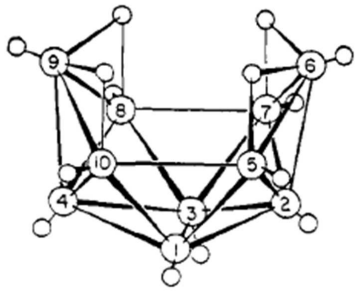
$$2H^+ + 6H_2O + [B_{20}H_{18}]^{2-} \longrightarrow B_{18}H_{22} + 2B(OH)_3 + 2H_2$$

The two $B_{18}H_{22}$ hydrides are structurally related to *nido*-decaborane(14) and may be formally generated by fusion of two decaborane-like fragments with sharing of two common (5-6) atoms in the fused structure. In Fig. 1.6 and 1.7, the structures [40, 41, 42, 43] and numbering systems of *anti*- and *syn*-$B_{18}H_{22}$ as well as $B_{10}H_{14}$ are shown. In the centrosymmetric *anti*-$B_{18}H_{22}$, the 5- and 6-boron atoms of one $B_{10}H_{14}$ fragment are, respectively, the 6- and 5-boron atoms of the second fragment. In the case of *syn*-$B_{18}H_{22}$ (C_2 symmetry) the 5- and 6-boron atoms bear the same relationship to both $B_{10}H_{14}$ fragments.



anti-$B_{18}H_{22}$ *syn*-$B_{18}H_{22}$

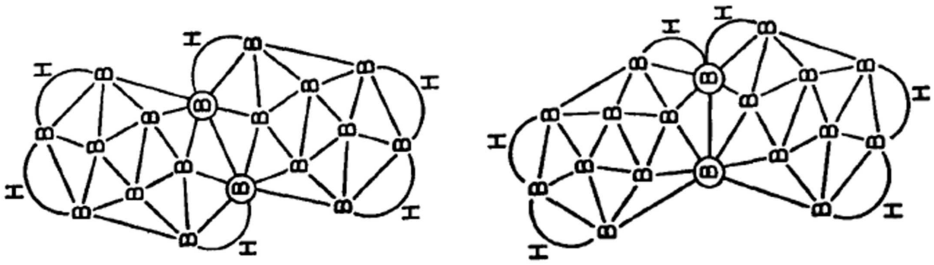
Figure 1.6: Structure and numbering for the *anti*-$B_{18}H_{22}$ and *syn*-$B_{18}H_{22}$.



$B_{10}H_{14}$

Figure 1.7: Structure and numbering for the $B_{10}H_{14}$.

A structural feature of $B_{18}H_{22}$ hydrides is the coordination of both B atoms of the shared (5-6) vertex to six other B atoms. In Fig. 1.8, a bridge H atom serves to connect each of the six-coordinated B atoms which do not have terminal H atoms. Both B atoms are encircled in the figure shown below.



anti-$B_{18}H_{22}$ *syn*-$B_{18}H_{22}$

Figure 1.8: Lipscomb's structures of *anti*-$B_{18}H_{22}$ and *syn*-$B_{18}H_{22}$ [42].

The chemistry of both isomers *anti*- and *syn*-$B_{18}H_{22}$ is diverse as studied by Hawthorne and co-workers [44]. Regarding acid-base reactions, titration of both isomers in aqueous media indicates that each isomer behaves as a diprotic acid. The first ionization constant of each isomer can be compared with the ionization of strong acids. The measurement of the second ionization constant

gives pK_a values of 7.5 and 8.6 for *anti*- and *syn*-$B_{18}H_{22}$, respectively. Unlike $B_{10}H_{14}$, the $B_{18}H_{22}$ hydrides do not have evidence of derivative formation in the presence of nucleophilic reagents such as acetonitrile. Decaborane(14) acquires two electrons when treated with sodium amalgam in tetrahydrofuran to produce $[B_{10}H_{14}]^{2-}$ ions. In analogous reduction reactions, *anti*- and *syn*-$B_{18}H_{22}$ react with sodium amalgam in tetrahydrofuran to produce the corresponding $[B_{18}H_{22}]^{2-}$ ions. When *anti*- and *syn*-$B_{18}H_{22}$ are subjected to base-catalyzed exchange with DCl in dioxane solution, only six deuterium atoms per molecule are incorporated. In both isomers, only two equivalent terminal BH groups are exchanged and the remaining four deuterium atoms enters bridge-hydrogen positions.

One of the most common applications of $B_{18}H_{22}$ hydrides is related to semiconductor manufacture [45, 46]. Large boron hydrides have become important feed stocks for boron doped p-type impurity regions in semiconductors. In particular, high molecular weight boron hydrides comprising at least a five boron atom cluster are preferred as boron atom feed stocks for molecular boron implantation.

An important aspect of modern semiconductor technology is the continuous development of smaller and faster devices, a process known as *scaling*. The greatest impact of scaling on ion implantation processes is the scaling of junction depths, which requires increasingly shallow junctions as the device dimensions are decreased.

There may be boron hydrides with the same number of boron and hydrogen atoms having different chemical properties, e.g. structural isomers or stereoisomers, and boron hydrides having the same number of boron atoms but different number of hydrogens. These boron hydrides, including isomers and compounds containing various amounts of hydrogen, are known as n-boron atom boranes. Mixtures of isomers and mixtures of n-boron atom boranes are suitable for use in the implantation methods.

1.4. Photochemistry of the DNA base pairs

The structure of DNA as a double helix is maintained by two types of interactions, hydrogen bonding and stacking. The hydrogen bonds orchestrate the base pairing between the Adenine (A) and Thymine (T) and the Guanine (G) and Cytosine (C) nucleic acids in the Watson-Crick DNA double helix [47, 48, 49]. Stacking interaction primarily determines the helix architecture [50]. The nucleic acid base pairs AT and GC of DNA referring to the Watson-Crick (WC) canonical ones are depicted in Fig. 1.9. The canonical AT base pair is composed between A and T due to two hydrogen bonds: O(T)$\cdots$H-N(A) and N(T)-H$\cdots$N(A), while the GC base pair is formed by G and C linked via three hydrogen bonds: N(C)-H$\cdots$O(G), N(C)$\cdots$H-N(G), and O(C)$\cdots$H-N(G) where the base is indicated within parentheses.

The role of the hydrogen bonds in DNA is twofold: 1) they are responsible for the electronic complementarity of the bases and 2) they hold the DNA shape complementarity along the double helix. Geometrically, the latter means that the distance R between H_9(purine) and H_1'(pyrimidine) falls within $10.6-10.8$ Å and glycosyl bond angles $\alpha_1 \equiv \angle N_9H_9$(purine)$H_1'$(pyrimidine) and $\alpha_2 \equiv \angle N_1'H_1'$(pyrimidine)$H_9$(purine) belong to the angle range $[52^o, 70^o]$ [51, 52]. The hydrogen-bond patterns corresponding to the complementary base pairs predetermine a genetic code alphabet within the Watson-Crick model of DNA [47, 48, 49, 53].

Altered hydrogen-bond patterns or non-Watson-Crick base pairings (mismatches or mispairings) might cause an alteration of the genetic code leading to point mutations [48]. A couple of typical altered hydrogen-bond patterns are schematically shown in Fig. 1.10. These patterns correspond to pairing of *rare* tautomers of the DNA bases which, as believed, cause the formation of mismatches producing spontaneous mutations in the genetic code [48].

In a pioneering work in 1963, P.-O. Lödwin [54, 55] introduced the hypothesis that the double proton transfer reaction in DNA base pairs could be a possible source of spontaneous mutations, since rare tautomers could be formed, which might disturb the genetic code. Since then, many theoretical studies have been carried out to check the Löwdin's hypothesis. The earlier studies [56, 57, 58, 59, 60, 61, 62], considered the double proton transfer process on the ground state of neutral pairs. Because of the size of the AT and GC, low computational levels of theory were used. These works used semiempirical

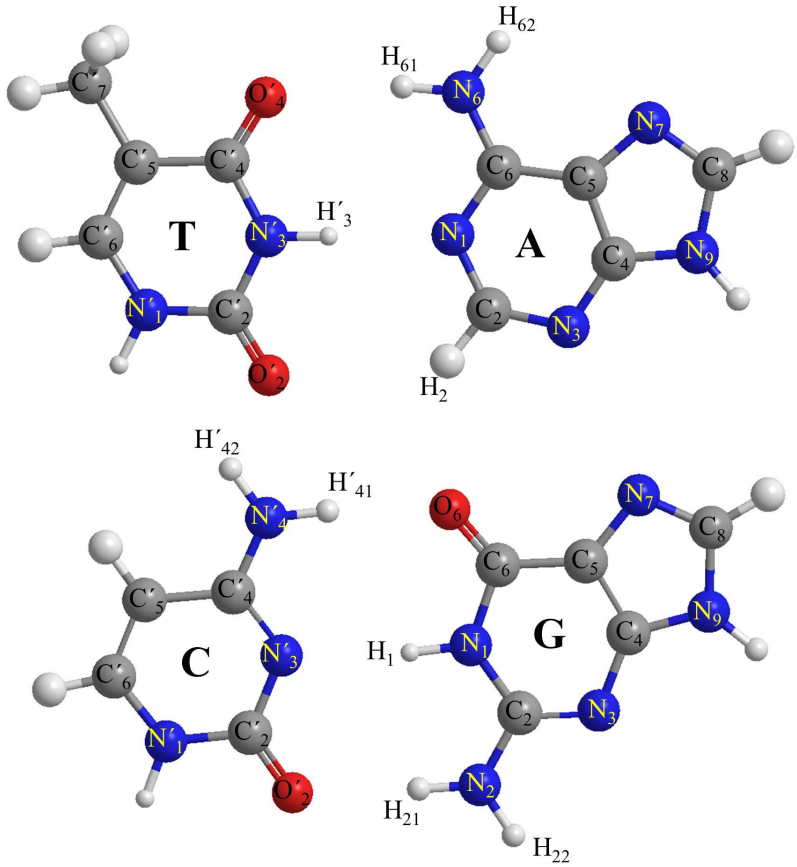


Figure 1.9: The canonical AT and GC base pairs of DNA.

[56, 57, 58, 59] and *ab initio* [60, 61, 62] methods to study the proton transfer processes at fixed geometries for the monomers. Consequently, both single and double proton transfer reactions were found to be unfavorable. Geometries were fully optimized later, finding smaller barriers [63, 64, 65]. All studies agreed that the single proton transfer (SPT) reaction was less favorable than the double proton transfer (DPT) one, because the SPT process implies a charge separation in the formation of the resulting ion-pair complex, while in the DPT reaction the neutrality is preserved. Nevertheless, the computed energy barrier of the DPT process was still high [65].

In the nineties, the DPT process induced by photoexcitation in DNA base pairs or model systems was of great interest, both from an experimental [66, 67, 68, 69, 70, 71, 72, 73, 74] and a theoretical point of view [75, 76, 77, 78]. In particular, Zewail et al. [66, 67] studied the double proton transfer reaction in the 7-azaindole dimer as a model system of DNA base pairs by means of femtosecond dynamic experiments. Other authors have also considered the same system by using similar techniques [68, 69, 70, 71, 72, 73, 74]. The main issues of discussion have been whether the mechanism of the reaction is stepwise or

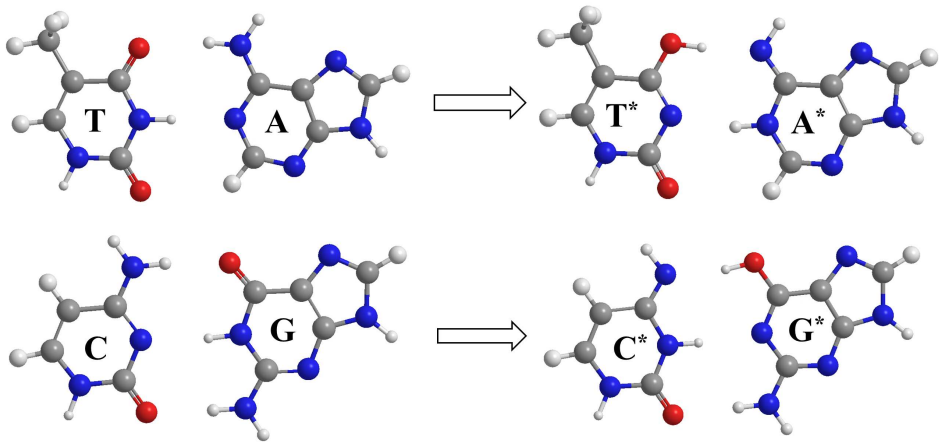


Figure 1.10: The double proton transfers in the canonical AT and GC base pairs of DNA.

concerted and the nature of the formed intermediate. Some theoretical studies [75, 76] on the lowest excited state of this system, using the CIS method, have shown that the DPT process is a two-step reaction with an important energy barrier, in contrast to what was obtained by Catalán et al. using fixed geometries derived from ground state optimizations [77]. Results from Ref. [76] show that the intermediate corresponding to the stepwise mechanism has a zwitterionic character. In addition, calculations provide another much more stable minimum in the region of the intermediate with a neutral character, since the transfer of a single proton is compensated by a charge transfer electronic excitation. Such intermediate could be responsible for the stepwise reaction observed in experiments through a nonadiabatic transition from the initially accessed electronic state to the one of the neutral intermediate [76]. In fact, a more recent theoretical study [79] on 7-azaindole has found two static stepwise mechanisms based on the presence of a neutral and an ionic intermediate after a proton being transferred from one moiety to the other. In a theoretical study [78] for the Guanine-Cytosine and Adenine-Thymine Watson-Crick base pairs, the computed energy barrier of the DPT reaction in the lowest singlet excited state turns out to be significant, whereas a charge transfer excited electronic state has been estimated at low energies in the region corresponding to a SPT that neutralizes the charge separation induced by the electronic transition. The state can be accessible through internal conversion from the initially populated electronic excited state. Domcke and co-workers [80, 81, 82] have characterized a conical intersection in the GC and AT base pairs between the charge transfer state and the ground state after the SPT reaction triggered by the charge separation of the charge transfer state. Thus, the DPT reaction completes the process of recovering the ground state of the canonical base pairs after being photoinduced. In the pyridine-pyrrole system, Frutos et al. [83] have found an avoided crossing between the lowest singlet excited state and the charge transfer state, at the CASSCF/CASPT2 level, which connects both states nonadiabatically. The latter theoretical studies [80, 81, 82, 83] explain the photostability of the Watson-Crick base pairs which are supported by experimental works in the femtosecond timescale [84, 85, 86, 87]. Nonetheless, nowadays, the debate on the hydrogen-bonded dynamics to explain the fast deactivation in photoinduced DNA is still open [86, 88, 89, 90, 91, 92].

Briefly, some experiments evidence that the proton or hydrogen transfer is not a relevant process in the photostability of DNA, in favor to other deactivation mechanisms. Furthermore, the effect of the environment points out the importance of the surrounding media (DNA or water solution) on the way the deactivation processes found *in vacuo* are altered.

Regarding tautomeric mispairing in the GC base pair, Florián and Leszczynsky [65] determined, at the HF/MP2 level, that the lowest-energy base pairs, among those resulting from the different arrangements of the protons shared in hydrogen bonds, are the canonical base pair and the rare tautomeric form imino-keto/amino-enol G^*C^*. The latter is formed by three hydrogen bonds: $N'_4(C)\cdots H\text{-}O_6(G)$, $N'_3(C)\text{-}H\cdots N_1(G)$, and $O'_2(C)\cdots H\text{-}N_2(G)$ which, in comparison to the canonical form, differ in the position of the protons in two hydrogen bonds: $N'_4(C)\text{-}H\cdots O_6(G)$ and $N'_3(C)\cdots H\text{-}N_1(G)$. Two different energetic pathways can govern the proton rearrangements to link the WC GC and G^*C^* base pairs. The higher-energy pathway is mediated by the mismatched ionic base pair 3H-amino-keto/amino-keto GC_{ionic} consisting of two transition structures GC_1^{ts} and GC_2^{ts} which link GC to GC_{ionic}, and GC_{ionic} to G^*C^*, respectively. In Fig. 1.11, the ground-state potential energy surface (PES) of the proton transfer in the GC base pair is depicted.

The lower-energy pathway passes through the transition pair-linker GC_3^{ts} which govern the DPT along $N'_4(C)\cdots H\cdots O_6(G)$ and $N'_3(C)\cdots H\cdots N_1(G)$ hydrogen bonds. The most significant geometrical changes are related to the bond length of the hydrogen bonds which are shrinked considerably. At the level of calculation of Ref. [65], basically the bond length $r(N'_3(C)\text{-}N_1(G))$ changes from 3.043 to 2.608 Å, $r(N'_4(C)\text{-}O_6(G))$ is contracted from 2.929 to 2.442 Å, and the $r(O'_2(C)\cdots H\text{-}N_2(G))$, which does not participate in the DPT, shows a certain decrease from 3.018 to 2.867 Å.

The mechanism proposed by Kryachko and Sabin [94] to explain the origin of mispairings in the AT base pair consists of four principal stages. Initially, the proton migrates along the weaker hydrogen bond $N'_3(T)\text{-}H\cdots N_1(A)$ of AT.

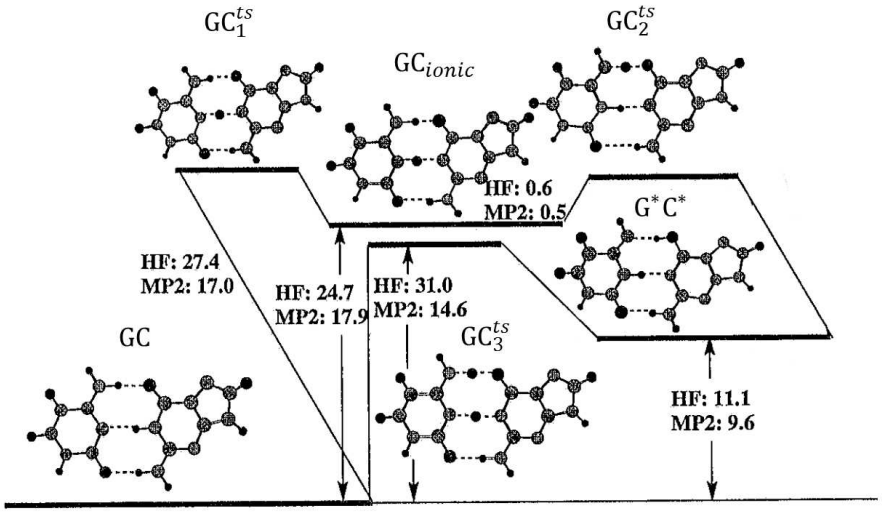


Figure 1.11: The proton transfer PES of the GC base pair. Energies are given in kcal·mol^{-1}; HF corresponds to HF/6-31+G(d), while MP2 refers to the MP2//HF/6-31+G(d). Ref. [93].

The transition structure AT_1^{ts} has to be surmounted to arrive at the ionic pair AT_{ionic} which is formed by two alligned hydrogen bonds, $N_3'(T)\cdots H\text{-}N_1(A)$ and $O_4'(T)\cdots H\text{-}N_6(A)$. This ionic structure results to be more stable than A^*T^* at the HF level. The existence of another transition state AT_3^{ts} in the vicinity of the AT_{ionic} strongly affects its metastability. This transition state is assigned to the relative motion of A and T within their pair and constitutes the source structure to the mispairing valley mechanism. In the second stage, the proton in AT_{ionic} surmounts the transition state AT_2^{ts} via migration along the $O_4'(T)\cdots H\cdots N_6(A)$ hydrogen bond. The two-stage process leads to the rare tautomeric pair A^*T^*. In Fig. 1.12, the ground-state potential energy surface of the proton transfer in the AT base pair is represented.

The mispairing valley (Fig. 1.13) explains the reaction pathway that covers the third and fourth stages of the mispairing mechanism in AT. It starts at

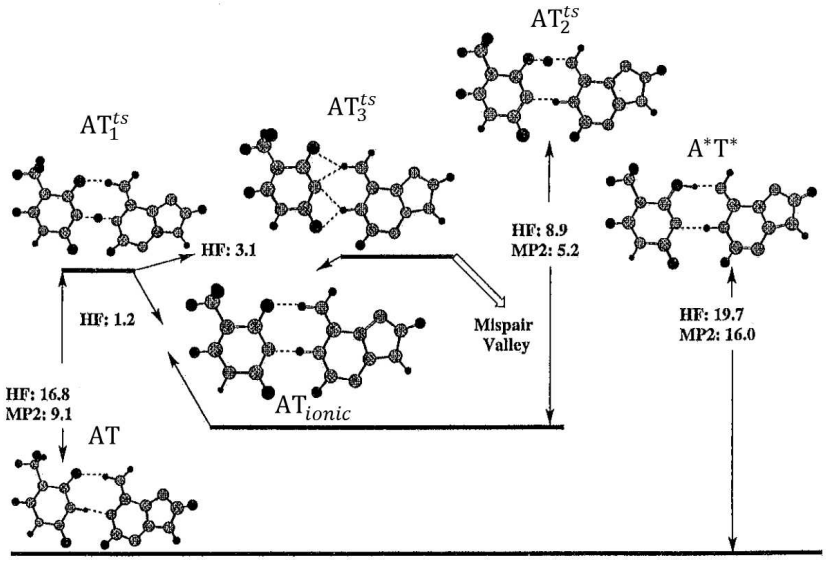


Figure 1.12: The portion of the PES of the AT base pair. Energies are given in kcal·mol^{-1}; HF corresponds to HF/6-31+G(d), while MP2 refers to the MP2//HF/6-31+G(d). Ref. [93].

the AT_3^{ts} structure which results in an interbase shift where the N_6(A)-H bond shares its proton with O_4' and N_3' atoms of thymine and N_1(A)-H bond forms a bifurcated hydrogen bond with N_3'(T) and O_2'(T). The proton transfer pathway ends at the energy minimum structure AT_3 concluding the third stage. A second proton is transferred in the fourth stage to connect AT_3 of ionic character with the neutral AT_4 or AT_5 mismatched pairs through the transition states AT_4^{ts} and AT_5^{ts}, respectively. Both activation barriers are low in energy but AT_4 is more stable than AT_5 at the HF level. So it is more favorable the ionic-to-neutral transition $AT_3 \rightarrow AT_4$ than $AT_3 \rightarrow AT_5$. The AT_4 structure is slightly more compact than the WC pair being characterized by $R = 10.023$ Å, and $\alpha_1 = 41.1^{\circ}$ and $\alpha_2 = 71.1^{\circ}$. The concerted route, unlike in GC base pair, is much higher in energy and, therefore, does not play an important role in the mispairing mechanism of AT.

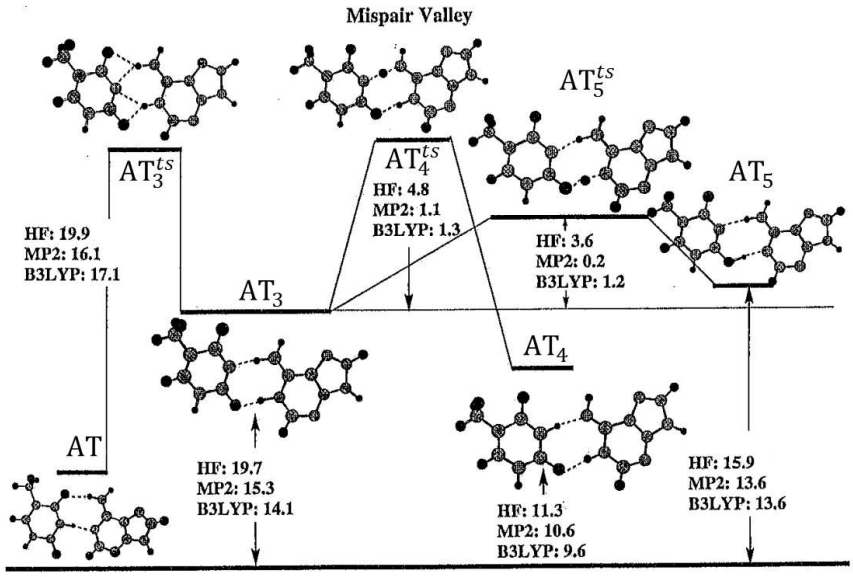


Figure 1.13: The mispairing valley in the PES of the AT base pair. Energies are given in kcal·mol^{-1}; HF corresponds to HF/6-31+G(d) while B3LYP refers to B3LYP/6-31+G(d) and MP2 denotes MP2//HF/6-31+G(d). Ref. [93].

In conclusion, two mechanisms of the tautomeric thermal conversions in GC and AT are presented schematically:

$$GC \rightarrow GC_3^{ts} \rightarrow G^*C^*$$

and

$$AT \rightarrow AT_1^{ts} \rightarrow AT_{ionic} \rightarrow AT_3^{ts} \rightarrow AT_3 \rightarrow AT_4^{ts} \rightarrow AT_4$$

In contrast to the single-step concerted quantum tunneling of two protons in the GC base pair, the mechanism of the hydrogen-bonded mispairing in the AT is sequential and includes at least three stages.

In the present thesis, the mechanisms of hydrogen and proton transfer are studied in the excited state using CASPT2//CASSCF methodology to explain both photostability and tautomerism in the DNA base pairs. This novel contribution might be important for the interpretation of previous experimental findings and the design of new experiments to unveil the role of hydrogen transfer in the excited state dynamics of DNA.

Chapter 2

Objectives

Several works have been carried out and discussed in the present thesis. The use of high-level quantum-chemical methods in the study of different systems after being irradiated with UV radiation is the most remarkable feature. The interest of such molecular arrangements are focus on the determination of the absorption spectra and the analysis of the mechanisms which explain the photophysics and photochemistry of the systems introduced in the previous chapter.

First, the multiconfigurational second-order perturbation theory restrited active space (RASPT2) method has been tested in the computation of electronic excited states of organic and inorganic molecules. RASPT2 calculations are compared with the well-established CASPT2 method or CCSD and experimental values when available. The main goal is to select systematic procedures in order to provide the most accurate results at reasonable computational costs and extend the use of the method to larger systems. One advantage of RASPT2 versus CASPT2 is the employment of large active spaces which allow the treatment of extended π systems like free base porphirin, or the inclusion of both valence and Rydberg states and molecular orbitals in computations

on molecules such as ethene and benzene. For first-row transition metals, the main aim is to analyze the effect on excitation energies of moving the $4d$ correlating shell from RAS2 to RAS3.

The RASPT2 method has also been used to compute the potential energy surfaces of stretched ethene to parameterize the extended Hubbard model which is employed to calculate the structure and excitation energies of conjugated and aromatic hydrocarbons. Bond lengths derived from the model are compared with MP2 and experimental values, and excitation energies obtained with the model are tested with RASPT2 calculations.

Experimental and theoretical work on boron hydrides has been one of the subjects in the present thesis. The distinct photophysics of the two isomers of octadecaborane(22), *anti*- and *syn*-$B_{18}H_{22}$, has been explained through the different rearrangements at their respective conical intersections. The production of singlet oxygen is also an interesting aspect in derivatives of these boron hydrides.

And finally, the proton/hydrogen transfer processes have been studied in the Adenine-Thymine and Guanine-Cytosine base pairs. The mechanisms described, together with the decay paths of isolated nucleobases and the formation of π-stacked pyrimidines molecules, explain the photostability and photoreactivity of DNA *in vacuo*. Since the biological environment might be important to evaluate such processes, QM/MM calculations are also carried out for a sample of DNA, [dG_9·dC_9], surrounded by water molecules. Furthermore, some dynamical insights are considered *in vacuo* to estimate which proton/hydrogen process dominates in DNA dynamics.

Chapter 3

Quantum-Chemical Methodology

Quantum chemistry is based on finding approximate solutions of the time independent Schrödinger equation (Eq. 3.1):

$$\hat{H}\Phi = E\Phi \tag{3.1}$$

Typically in the non-relativistic formulation framework and within the Born-Oppenheimer approximation [95, 96, 97, 98]. The Hamiltonian operator for a system with N electrons and M nuclei is defined in atomic units by the following expression:

$$\hat{H} = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{N}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} \tag{3.2}$$

The first term accounts for the kinetic energy of the electrons, the second one expresses the Coulomb attraction between electrons and nuclei, the next term represents the Coulomb repulsion between the electrons, and finally a constant nuclear repulsion for a given geometry is included. The last term in Eq. 3.2 adds a constant to the energy eigenvalues and has no effect on the Hamiltonian eigenfunctions. The constant term is omitted to obtain the electronic Hamiltonian, $\hat{H}_{el}$. Solving the electronic Schrödinger equation, $\hat{H}_{el}\Psi = E_e\Psi$, for different positions of the nuclei, a potential energy hypersurface (PEH) on which the nuclei move is obtained. The electronic wave function depends explicitly on the electron coordinates but parametrically on the nuclear variables and the total energy for fixed nuclei must also include the constant nuclear repulsion V_N as shown below:

$$\Psi = \Psi\left(\{\vec{r}_i\};\{\vec{R}_A\}\right) \qquad E(\vec{R}_A) = E_e(\vec{R}_A) + V_N \tag{3.3}$$

The antisymmetry principle states that the electronic wave function must be antisymmetric with respect to the interchange of the coordinates (both space and spin) of any two electrons [95]. The major difficulty to solve the electronic Schrödinger equation is how to treat the two-electron term $\frac{1}{r_{ij}}$ which takes into account the non-independent movement of the electrons (electronic correlation). Each electron is surrounded by a *Coulomb hole* to account for the null probability of finding two electrons in the same infinitesimal volume element. The correlation among the electrons is not treated properly, however, by many approximate methods.

The simplest wave function to describe a many-electron system is a Slater determinant, Ψ_{SD}, consisting of orthonormal one-electron spin orbitals, $\chi_i(\vec{r}_j,\omega)$, which are built from the composition of spatial orbitals, $\Psi_i(\vec{r})$, and spin functions, $\alpha(\omega)$ or $\beta(\omega)$. Such wave functions introduce exchange correlation, which means that the motion of two electrons with parallel spins is correlated (*Fermi hole*). This is a consequence of the antisymmetry condition. Nevertheless, there is a probability different from zero to find two electrons with

different spin at the same point in the space. This means that the Slater determinant is not correlated. For a N-electron system, the Slater determinant is expressed as follows:

$$\Psi_{SD}(\mathbf{x}_1, \mathbf{x}_2 \cdots \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_i(\mathbf{x}_1) & \chi_j(\mathbf{x}_1) & \cdots & \chi_k(\mathbf{x}_1) \\ \chi_i(\mathbf{x}_2) & \chi_j(\mathbf{x}_2) & \cdots & \chi_k(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_i(\mathbf{x}_N) & \chi_j(\mathbf{x}_N) & \cdots & \chi_k(\mathbf{x}_N) \end{vmatrix} \tag{3.4}$$

with $\chi_j(\mathbf{x}_i)$ the molecular spin orbitals and $\mathbf{x}_i$ the spatial $\vec{r}$ and spin ω coordinates of electron i. Although the Schrödinger equation looks rather simple, it is impossible to solve exactly except for e.g. one-electron, mono- and diatomic molecules. Therefore, the main challenge for computational chemistry is to find approximations to the exact solution of the Schrödinger equation.

There are two types of quantum methods widely used today:

1. *Ab initio methods* [95, 99] (see Fig. 3.1), based on solving the Schrödinger equation for the whole molecular system without any parameterization.

2. *Parameterized methods*, including the classical semiempirical approaches, where the parameterization is focused on avoiding the calculation of the two-electron integrals, and the modern semiempirical Density Functional Theory, where electronic density is used instead of the wave function for the energy calculations. For the latter methods, the analytical expression between energy and electronic density is unknown, therefore approximate relations including some parameters are needed. As it occurred in the classical semiempirical approaches, DFT methods often have erratic and unexpected failures and might work out well or not depending on the molecular system under study and the particular electronic structure problem. This behavior make them not very systematic, difficult to calibrate, and in some cases without predictive character. They will not be discussed in the present work.

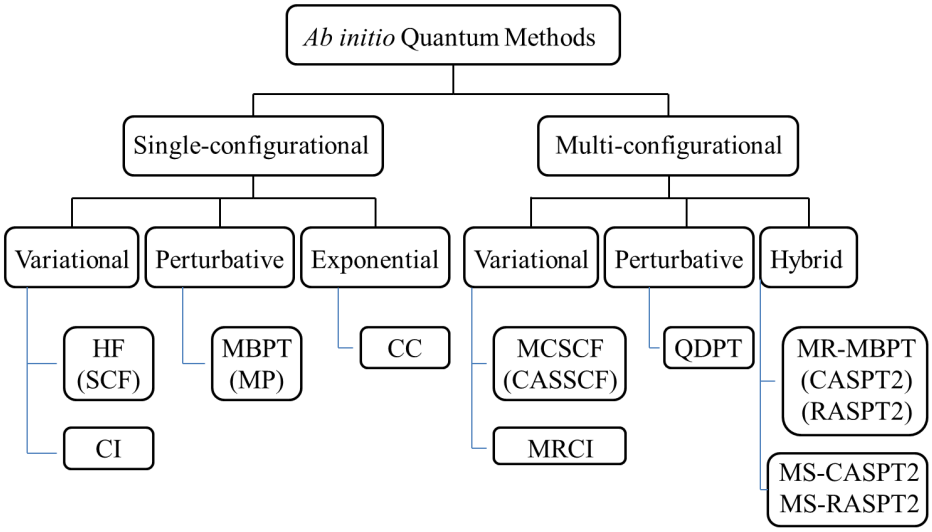


Figure 3.1: Scheme of the most common quantum-chemical *ab initio* methods.

3.1. The Hartree-Fock Method

The Hartree-Fock (HF) approximation [100, 101] usually constitutes the starting point of almost all more accurate methods. The HF procedure provides the mathematical tools to obtain the unknown spin orbitals to build the best Slater determinant by making use of the variational principle [95]. The spin orbitals verify the canonical equation 3.5 derived from the stationary condition $\delta\langle\Psi|\hat{H}_{el}|\Psi\rangle = 0$ to minimize the electronic energy $E_e = \frac{\langle\Psi|\hat{H}_{el}|\Psi\rangle}{\langle\Psi|\Psi\rangle}$ over arbitrary variations of the spin orbitals:

$$f(i)\chi(\mathbf{x}_i) = \varepsilon\chi(\mathbf{x}_i) \tag{3.5}$$

where $f(i)$ is an effective one-electron operator, called the Fock operator, of the form:

$$f(i) = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + v^{HF}(i) \tag{3.6}$$

The Hartree-Fock potential for *ith* electron depends on the spin orbitals of the other electrons. The essence of the HF approximation is to replace the complicated many-electron problem by a one-electron problem in which electron-electron repulsion is treated in an average way. As energy depends on spin orbitals, the HF equation is not linear and must be solved iteratively in a self-consistent field (SCF) approach [95].

The HF equation can be solved numerically for atoms when the wave function is built from spin orbitals. In the molecular case, an analytical expression for the spatial orbitals is needed. Usually, the Linear Combination of Atomic Orbitals is employed (LCAO):

$$\psi_i(\vec{r}) = \sum_{\mu=1}^{K} c_{\mu i}\phi_\mu(\vec{r}) \tag{3.7}$$

which are substituted in the HF equation to give as result the Roothaan equations [95, 102]. In turn, atomic orbitals are expanded on basis functions which span the space to describe the electron distribution over the whole molecule. For an infinite number of basis functions, one reaches the Hartree-Fock limit, E_0.

The Hartree-Fock approximation does not provide an accurate determination of atomic or molecular properties, mainly because it does not take into account the electron correlation energy. The correlation energy, E_{corr}, is defined as the

difference between the exact non-relativistic energy of the system within the Born-Oppenheimer approximation, ε_0, and the Hartree-Fock limit, E_0:

$$E_{corr} = \varepsilon_0 - E_0 \tag{3.8}$$

Since ε_0 is lower than E_0, the correlation energy is negative. As the use of a complete basis set is not feasible, the exact electron correlation of a system cannot be computed, except for small systems. E_{corr} must be obtained thorough approximate treatments which use finite basis sets. Such methods are known as post-Hartree-Fock procedures. The HF method determines the best single determinant trial wave function, Φ_{HF}, within a given one-electron basis set. In order to improve the HF results, the starting point must be to build a trial wave function which contains more than one Slater determinant. A generic multi-determinant wave function can be written as:

$$\Psi = a_0\Phi_{HF} + \sum_{i=1} a_i\Phi_i \tag{3.9}$$

where $|a_i|^2$ is the weight of *ith* determinant in the expansion. The many-electron bases, Φ_i, used to span the wave function, are constructed by replacing molecular orbitals (MOs), or spin orbitals, which are occupied in the HF determinant by MOs which are unoccupied. These Φ_i are denoted according to how many occupied HF MOs have been replaced by unoccupied MOs, i.e. Single (S), Double (D), Triple (T), etc. The truncation of a set of Slater determinants limits the description of the correlation energy, and therefore, the quality of calculations. The difference among the many post Hartree-Fock methods is how the correlation energy is computed. According to this, the most usual *ab initio* approaches are: Configuration Interaction (CI), Many Body Perturbation Theory (MBPT), and Coupled Cluster (CC).

3.2. Configuration Interaction (CI)

Configuration Interaction [103] is a method based on the variational principle that uses as a trial function, ψ_{CI}, one expanded by a linear combination of Slater determinants. These determinants are built by MOs taken, for instance, from a Hartree-Fock calculation and which are kept fixed during the convergence process:

$$\psi_{CI} = c_0|\psi_0\rangle + c_S|\psi_S\rangle + c_D|\psi_D\rangle + c_T|\psi_T\rangle + c_Q|\psi_Q\rangle + \cdots \tag{3.10}$$

In this method, the expansion coefficients, c_i, in equation 3.10, or a_i in the general expression 3.9, are determined by applying the stationary condition to the energy. If all possible excited determinants (singly, doubly, etc.) are included in the CI function, the method is called *full CI* and it yields the exact solution within the one-electron basis set employed. In practice, because of computational restrictions, the number of excited determinants in the CI expansion must be reduced, leading to different methods as CID (including doubly excited determinants), CISD, CISDT, etc. One problem of truncated CI methods is that they are not size extensive, i.e. those methods do not scale properly with the number of particles. This lack of size extensivity is the reason why CISD retrieves less and less percentage of electron correlation as the systems are enlarged. Additionally, when the HF wave function is selected as reference, the methods cannot handle many electronic structure problems of complex nature which require a multireference description.

The Multi-Configurational Self-Consistent Field (MCSCF) wave function can be considered as an extension of the CI wave function [97, 98] where the coefficients of the expansion and the MOs are both optimized under the variational principle. Such wave function includes therefore several to many CIs which can be used as reference for more elaborated multiconfigurational post-HF calculations. The well-known Complete Active Space SCF (CASSCF) method

[1, 104, 105] is a variant of the MCSCF treatment that has become popular because of its technical and conceptual simplicity. Here, the selection of the CIs forming the multireferential space is performed by including all those MOs and electrons which are relevant in the chemical process studied. Thus, MOs are partitioned into three subspaces: active - variable occupation -, inactive - always doubly occupied -, and external - always unoccupied in the CASSCF wave function. Within the active subspace a full CI (CAS CI) is performed and all the proper symmetry- and spin-adapted configuration state functions, CSFs, are included in the MCSCF optimization. The goal of the CASSCF method is to obtain an improved wave function before proceeding to calculate the remaining part of correlation energy. There are two types of correlation, non-dynamical and dynamical correlation. The former is associated with the error caused by the HF approximation in cases of near-degeneracy between different electronic configurations, while the latter is related to the fact that the movement of the electrons depends on the instantaneous positions instead of the average field of the other electrons. The valence CASSCF method recovers the non-dynamical correlation and gives a multireference wave function used as reference in more accurate procedures which retrieve the dynamical correlation.

The Multi-Reference Configuration Interaction (MRCI) approach [106, 107] normally uses a MCSCF wave function as reference for the subsequent CI calculation. In this case, a CISD would involve excitations of one or two electrons out of all the determinants which constitute the MCSCF wave function. The advantages of such methods are that they can generate very accurate wave functions if the reference MCSCF wave function includes enough configurations. The drawbacks are their non size-extensivity and their computational cost, typically too expensive even when truncating the (MR)CI expansion at the one-electron and two-electron promotion level, which make the method non applicable for medium to large systems.

3.3. Many-Body Perturbation Theory (MBPT)

MBPT-like methods [95, 97] apply perturbation theory techniques to the calculation of the energy and the wave function, obtained as solutions of some *ab initio* approaches. All methods derived from MBPT are size extensive. The resulting energy and wave function can be written as a Taylor expansion in powers of the perturbation parameter, λ (Eq. 3.11). Then corrections can be at first-, second-, etc., order depending on the terms of the Taylor expansion considered:

$$\begin{aligned} E &= E_0 + \lambda E_1 + \lambda^2 E_2 + \cdots \\ |\psi\rangle &= |\psi_0\rangle + \lambda|\psi_1\rangle + \cdots \end{aligned} \tag{3.11}$$

More details on the basics of the theory are discussed in section 1.1.

Within these approaches, the most common are those based on the Møller-Plesset (MP) perturbation theory [108], which include MP2, MP3, etc., depending on the level of the correction. MP2 uses as reference a HF wave function and retrieves around 80-90% of the correlation energy [95]. Furthermore, among the *ab initio* methods for ground-state problems, it is the most inexpensive method to include electron correlation and it has been often used for geometry optimization, although nowadays has been substituted by the DFT approaches, in particular because the minor influence of the spin contamination problems in the latter for open-shell cases. In such situations, the reference wave function has to be of the unrestricted HF (UHF) type, and it is not an eigenfunction of the total spin operator. In practice, the ground-state solution tends to be contaminated by excited states and the perturbative treatment cannot correct the deficiency. Just as single reference CI can be extended to MRCI, it is also possible to use perturbation methods with a multi-determinant reference wave function leading to MR-MBPT. The most

popular is the CASPT2 method [9, 109] that is based on a CASSCF reference. In fact, CASPT2//CASSCF strategy is employed to compute excited states due to their good predictions (within 0.1 - 0.2 eV) [96] and a moderate computational cost.

3.4. Coupled Cluster (CC)

Perturbation methods add all types of corrections to the reference wave function to a given perturbative order. The idea in the Coupled-Cluster (CC) methods [110, 111, 112] is to include all corrections of a given type to infinite order [113]. Examples of them are CCSD, CCSDT, or CCSDTQ [95]. Usually, and for practical reasons, the most employed approaches truncate the equations to single- and double-type excitations, and in some cases to triple terms estimated by perturbation, CCSD(T). The cost of the procedure is extremely high, although the accuracy is quite good for well-defined closed-shell ground-state situations. As in MP methods, CC methods are size extensive. Standard CC theory is based on the exponential ansatz and takes the HF wave function as reference (Eq. 3.12):

$$|\psi_{CC}\rangle = exp(\hat{T})|\psi_{HF}\rangle = \sum_{k=0}^{\infty} \frac{1}{k!}\hat{T}^k|\psi_{HF}\rangle \qquad \hat{T} = \sum_{i=1}^{N} \hat{T}_i \tag{3.12}$$

where N is the number of electrons and $\hat{T}_i$ operator generates all ith excited Slater determinants from the HF reference.

As mentioned, these methods are computationally very demanding. CI methods are less used nowadays, to the profit of MP and, especially, CC methods. It is now proved that size extensivity is important for obtaining accurate results. In terms of accuracy, with a medium-sized basis set, the following order is often observed for ground-state calculations [99]:

$$\text{HF} << \text{MP2} < \text{CISD} < \text{CCSD} < \text{MP4} < \text{CCSD(T)}$$

All of these methods are, however, single-reference-based approaches. There is a large number of electronic structure situations in which they cannot be properly applied due to a poor description of the HF reference. For instance, biradical situations, degeneracy cases, most of bond dissociations, electronic excited states, heavy elements with many degenerate states, etc. In order to avoid these problems the immediate solution is the use of multireferential approaches. The multiconfigurational RASSCF/RASPT2 and CASSCF/CASPT2 methods have been used in the computation of the results exposed in the present thesis. The basics of these procedures are discussed in section 1.1.

3.5. Hybrid Methods (QM/MM)

Some systems with a high number of atoms (enzymes or proteins) cannot be treated by quantum-mechanical (QM) methods which are limited to medium-sized and small systems. In order to model such complex systems, methods based on Molecular Mechanics (MM) are not accurate to describe the changes of the electronic structure in the chemical process. Bond formation and bond breaking, charge transfer and/or electronic excitation require a QM treatment. Due to this demanding situations, a hybrid method (QM/MM) has been developed to study biological systems in which a combination of algorithms based on QM and MM are used to obtain the accuracy of QM with the low computational cost of MM.

The main idea of the hybrid methods QM/MM [114, 115] consists of dividing the system into two regions, one of them is treated by QM and the other one by MM, as shown in Fig. 3.2. The total energy of the system can be defined as the sum of the energy contributions of the parts QM, MM, and the interaction between both of them:

$$E_{tot} = E_{QM} + E_{MM} + E_{QM/MM} \tag{3.13}$$

The atoms of the QM region are described by a Hamiltonian related to equation 3.2. The MM zone, which contains the other atoms of the system, will be treated classically and their interactions will be determined by empirical potential energy functions without taking into account the bond formation or bond breaking and the charge transfer phenomena [116]. The MM energy term is evaluated by force fields (FF) based on empirical potentials, small vibrational amplitudes, torsion, Van der Waals interactions, and electrostatic interactions. The MM energy depends initially on the atomic positions and the potentials of the nuclei when the polarization of the MM region is neglected. The $E_{QM/MM}$ term includes the interactions between the QM and MM atoms, i.e. the coupling between both regions of the system. The treatment of the short-range QM/MM interaction (Van der Waals) is based on a MM calculation. Regarding the electrostatic terms, there are several levels of approximation. According to Bakowies and Thiel [115], different schemes can be defined: *Mechanical Embedding* (ME), *Electrostatic Embedding* (EE), and *Polarized Embedding* (PE).

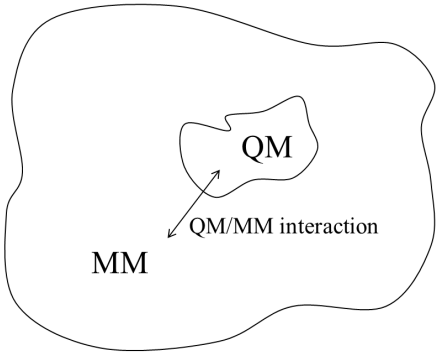


Figure 3.2: Scheme of the division of a molecular system into the QM and MM subsystems employed in the QM/MM approach.

In the ME scheme, the QM subsystem is treated computationally in the absence of the MM subsystem whereas the interaction between both regions is obtained at the MM level. In the EE treatment, the perturbation that the MM region generates on the QM zone is also considered in the QM calculation. Therefore, the polarization of the QM region, due to the charge distribution in the MM subsystem, is taken into account at the QM level. In the PE scheme, the polarization of the MM region, as result of the QM charge distribution, is also included in the QM/MM calculation. Despite both QM and MM subsystems are polarized, the change in the MM charge distribution is not propagated to the QM calculation and, therefore, the total energy is not variational. As an improvement of the PE scheme, the *Self-Consistent Embedding* scheme, SCE, has been developed to find the autoconsistency of the QM and MM polarizations via an iterative process.

Other aspect to consider in QM/MM calculations is the boundary region between the QM and MM subsystems. In many occasions, the division of the QM and MM regions is proceeded by means of covalent bonds as shown in Fig. 3.3. A treatment of the boundary region is required in order to satisfy the valences of the quantum atoms involved in the linkage of the quantum and classical regions. The most popular approximations are *Link Atom*, LA, *Local Self-Consistent Field*, LSCF, and *Generalized Hybrid Orbital*, GHO. In Fig. 3.3 a scheme of the LA model is shown and consists basically in using atoms not present in the original molecular system to link QM atoms in the boundary region. The most common link atom is hydrogen (HL). Link atoms are treated by QM methodology, so these additional QM centres in the model generates a variation in the coordinates and forces. Moreover, in the polarized QM/MM schemes (EE, PE, and SCE), overpolarization effects are obtained due to a fictitious polarization between Q1 and L atoms induced by the point charge of M1 atom.

The inclusion of new coordinates related to the link atoms gives rise additional degrees of freedom and, therefore, the QM/MM energy of the system is defined ambiguously. This drawback can be avoided when the L atom is placed in the line which connects Q1 and M1 and the L coordinates are expressed as a

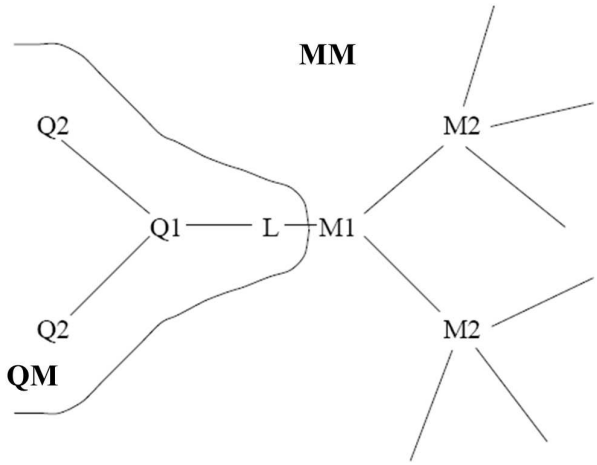


Figure 3.3: Labeling of atoms in the QM and MM regions employed in the QM/MM method when the boundary region breaks a covalent bond and the LA model is used. The boundary MM atom is labeled M1, whereas the MM atoms bounded to M1 are labeled M2, and so on. The QM atom bounded to M1 is named Q1 and the QM atoms linked to Q1 are named Q2, and so on. L corresponds to the link atom. Taken from Ref. [114]

function of the Q1 and M1 coordinates. The $R(Q1-L)$ distance can be scaled with respect to the $R(Q1-M1)$ distance by a scale factor, C_{L}, defined as:

$$R(\mathrm{Q1}-\mathrm{L}) = C_{\mathrm{L}} R(\mathrm{Q1}-\mathrm{M1}) \tag{3.14}$$

Thus, the $R(\mathrm{Q1}-\mathrm{L})$ and $R(\mathrm{Q1}-\mathrm{M1})$ equilibrium distances will have to satisfy the relationship 3.14 during the QM/MM geometry optimization. The scale factor depends on the nature of the broken bonds in the boundary region and the forming bonds when using the LA model. If C-C bonds are replaced by C-H bonds, the scale factor takes the value of 0.71 [117].

The overpolarization of the Q1−L bond, due to the proximity of the M1 point charge, can be solved by means of a variation in the distribution of the classical charge in the boundary region. Setting to zero the M1 charge in the QM Hamiltonian is the simplest way to avoid the overpolarization, but the net charge of the MM region can be changed and, therefore, the computed energy of the system may result to be inaccurate. A common solution to overpolarization consists in distributing the M1 charge, which is set to zero, among the M2 atoms. This approximation is valid when the charge portions are small. In the present thesis, the QM/MM calculations have been carried out by using the LA model and distributing the M1 charge among the M2 centres.

3.6. Nonadiabatic dynamics by the Surface-hopping method

Several mixed quantum-classical models have been developed to treat nonadiabatic quantum dynamics in a classical description [118]. The most popular approach to describe ultrafast molecular photoprocesses as electron transfer, internal conversion, and nonadiabatic photoisomerization is the surface-hopping method [119, 120, 121] in combination with an on-the-fly *ab initio* evaluation of the potential energy surfaces [122, 123, 124]. In the surface hopping (SH) method, classical trajectories are propagated on a single adiabatic potential energy surface until, according to some *hopping criterion*, a transition probability to another potential energy surface is computed and, depending on the comparison of the transition probability with a random number, the trajectory *hops* to the other adiabatic surface.

In the SH approach, atomic motions evolve classically on an effective potential given by the expectation value of the electronic Hamiltonian whereas a quantum-mechanical procedure is required in the computation of electronic transition probabilities. Excited-state energies, analytical energy gradients

with respect to the nuclear coordinates, and nonadiabatic coupling vectors are the electronic quantities needed to solve the Newton's equations for the nuclei and the electronic Schrödinger equation. As the effective potential that governs the nuclear motion is obtained from the electronic Hamiltonian, the electronic wave function is determined self-consistently with the trajectory. The direct or on-the-fly approach computes the potential energy surfaces and the nonadiabatic couplings at each time step during the numerical integration of the dynamical equations.

The Verlet algorithm is used in the numerical integration of the Newton's equations where the energy gradients are needed to compute the nuclear positions [125]. The fewest switches algorithm is the stochastic approach used to determine a possible hopping of the trajectory between adiabatic potential energy surfaces [118, 120].

- **The fewest switches algorithm**

The molecular system is described by the Hamiltonian $H = T_R + H_{el}(\mathbf{r};\mathbf{R})$ which accounts for the nuclear kinetic energy and the electronic Hamiltonian, dependent on the electronic coordinates ($\mathbf{r}$) and parametrically on the nuclear coordinates ($\mathbf{R}$). An orthonormal set of electronic basis functions $\phi_j(\mathbf{r};\mathbf{R})$ is chosen to define the matrix elements of the electronic Hamiltonian $V_{ij}(\mathbf{R}) = \langle\phi_i(\mathbf{r};\mathbf{R})|H_{el}(\mathbf{r};\mathbf{R})|\phi_j(\mathbf{r};\mathbf{R})\rangle$ and the nonadiabatic coupling vectors $\mathbf{d}_{ij}(\mathbf{R}) = \langle\phi_i(\mathbf{r};\mathbf{R})|\nabla_R\phi_j(\mathbf{r};\mathbf{R})\rangle$, where the gradient is defined with respect to the nuclear coordinates. The electronic wave function is expanded in terms of the electronic basis functions $\psi(\mathbf{r},\mathbf{R},t) = \sum_j c_j(t)e^{-iE_jt/\hbar}\phi_j(\mathbf{r};\mathbf{R})$. Substituting the wave function into the time dependent electronic Schrödinger equation, multiplying from the left by $\phi_k(\mathbf{r};\mathbf{R})$ and integrating over the electronic coordinates, it is obtained the following equation:

$$i\hbar\dot{c}_k = \sum_j c_j e^{i\nu_{kj}}(V_{kj} - i\hbar\dot{\mathbf{R}}\cdot\mathbf{d}_{kj}) \tag{3.15}$$

where $\nu_{kj} = \frac{1}{\hbar}\int_0^t (E_k - E_j)dt'$. Thus, excited-state energies and nonadiabatic couplings are needed to solve the time dependent Schrödinger equation.

In the adiabatic representation, the wave function is a linear combination of adiabatic electronic basis functions $\psi(\mathbf{r}, \mathbf{R}, t) = \sum_j a_j(t)\phi_j^{ad}(\mathbf{r}; \mathbf{R})$ and, as the matrix elements V_{ij} are vanished, the Schrödinger equation 3.15 is expressed as:

$$\dot{a}_k = -\sum_j a_j e^{i\nu_{kj}} \dot{\mathbf{R}} \cdot \mathbf{d}_{kj} = \sum_j A'_{kj} a_j \tag{3.16}$$

The equation 3.16 can be rewritten in a density-matrix notation as:

$$\dot{\rho}_{kk} = 2Re \sum_j A_{kj}\rho_{kj} \tag{3.17}$$

where $\rho_{kk} = a_k^*(t)a_k(t)$. The expression 3.17 describes the time evolution of electronic state populations. Introducing a time step Δt small enough, the change $\Delta\rho_{kk}$ of the diagonal matrix elements can be approximated by:

$$\Delta\rho_{kk} = \dot{\rho}_{kk}\Delta t \tag{3.18}$$

Moreover, the populations ρ_{kk} change according to the equation:

$$\Delta\rho_{kk} = -\sum_j p_{k\to j}\rho_{kj} + \sum_j p_{j\to k}\rho_{jk} \tag{3.19}$$

For a two-state system ($k, j = 1, 2$), combining the equations 3.17, 3.18, and 3.19, the electronic transition probabilities are derived as:

$$\begin{aligned} p_{1\to 2} &= 2ReA_{21}\Delta t\rho_{21}/\rho_{11} \\ p_{2\to 1} &= 2ReA_{12}\Delta t\rho_{12}/\rho_{22} \end{aligned} \tag{3.20}$$

where $p_{1\to 2}$ or $p_{2\to 1}$ is the transition probability when the system is initially in state 1 or state 2, respectively. Considering a large number N of trajectories, the number of trajectories assigned to state 1 at time t is $N_1' = \rho_{11}' N$. Similarly, $N_2' = \rho_{22}' N$ at time t. At a short time $t + \Delta t$ later, the state populations have changed slightly to ρ_{11} and ρ_{22}. Assuming $\rho_{11} < \rho_{11}'$ and $\rho_{22} > \rho_{22}'$, the number of trajectories switching from state 1 to state 2 in the interval of time Δt must exceed the number switching from 2 to 1. The fewest number of trajectory switches is accomplished when no switches from state 2 to 1 occur and $(\rho_{11}' - \rho_{11})N$ switches from state 1 to 2 occur. The probability to have a switch in the small interval of time Δt is expressed as:

$$\frac{\rho_{11}' - \rho_{11}}{\rho_{11}'} \approx \frac{\dot{\rho}_{22}\Delta t}{\rho_{11}} = p_{1\to 2} \tag{3.21}$$

The algorithm for the two-state system consists of the following steps:

Step 1. The trajectory is in state 1 at the integration step i.
Step 2. The trajectory is integrated one time interval Δt on state 1 to step $i + 1$. The equation 3.17 is integrated at step $i + 1$ to obtain the state populations ρ_{11} and ρ_{22}.
Step 3. A uniform random number ζ between 0 and 1 is generated and a trajectory switch from state 1 to 2 will be invoke if $p_{1\to 2} > \zeta$. In case the system is in state 2, a switch to state 1 will be invoked if $p_{2\to 1} > \zeta$.
Step 4. If no switch occurs (the vast majority of outcomes), return to step

2. If a switch from state 1 to 2 takes place, the trajectory will evolve on the potential energy surface of state 2. A velocity adjustment of the nuclear motion must be made to conserve the total energy. SH models do this adjustment to the component of velocity in the direction of the nonadiabatic coupling vector. After the velocity adjustment, return to step 2.

This procedure is repeated until the trajectory is considered to be finished and the sequence of steps is iterated for as many independent trajectories as required to obtain statistically significant conclusions. Quantities of interest derived from dynamics calculations such as quantum yields, state populations, or transient spectra are computed as avarages over all trajectories.

Chapter 4

Basis Sets

In *ab initio* methods, it is common to use Molecular Orbitals (MOs) to describe the electronic structure of a system. These MOs are obtained from a linear combination of atomic basis functions (LCAO-MO). The MO can be expressed as a function in the space generated by a complete basis set which implies the use of an infinite number of basis functions. In practice, a finite basis set is used instead. The dimension of the basis set and the precision of the *ab initio* computations have to be evaluated within the computational cost frame. There are two types of basis functions: Slater Type Orbitals (STO) [126] and Gaussian Type Orbitals (GTO) [127]. The Slater Type Orbitals are defined by the expression:

$$\chi_{\zeta,n,l,m}(r,\theta,\varphi) = N\, Y_{l.m}(\theta,\varphi) r^{n-1} e^{-\zeta r} \tag{4.1}$$

where ζ is the Slater orbital exponent, N is a normalization constant, $Y_{l,m}$ are the spherical harmonics, and r is the distance between the electrons and nuclei. This basis function is the exact solution of the Schrödinger equation for

the hydrogen atom. The STO-type orbitals do not have radial nodes which are introduced by a linear combination of different STOs. The exponential distance dependence ensures the rapid convergence with increasing number of basis functions. The Gaussian Type Orbitals can be expressed in spherical or cartesian coordinates:

$$\begin{aligned} \chi_{\alpha,n,l,m}(r,\theta,\varphi) &= N\, Y_{l.m}(\theta,\varphi) r^{2n-2-l} e^{-\alpha r^2} \\ \chi_{\alpha,l_x,l_y,l_z}(x,y,z) &= N x^{l_x} y^{l_y} x^{l_z} e^{-\alpha r^2} \end{aligned} \tag{4.2}$$

where α is the Gaussian orbital exponent and l_x, l_y, and l_z determine the orbital type (e.g., $l_x + l_y + l_z = 1$ is a p-orbital). The main difference between both formalisms is the simplicity when using the spherical functions. For instance, a d-type GTO has five spherical components ($Y_{2,-2}$, $Y_{2,-1}$, $Y_{2,0}$, $Y_{2,1}$, $Y_{2,2}$) but there are six components in cartesian coordinates (x^2, y^2, z^2, xy, xz, yz). These six functions can be combined to give the five d-type spherical functions and an additional s-type function ($x^2+y^2+z^2$). In a similar manner, the ten f-type functions in cartesian coordinates can be expressed in such a way to obtain seven f-type and three d-type spherical functions. The use of spherical functions for d and higher angular momentum basis functions in the determination of the two-electron integrals in an SCF calculation reduce the computing time considerably and avoid the linear dependence problem in extended basis sets.

The orbital exponents, ζ and α, determine the diffuseness of the basis functions. Thus, a large exponent implies a small dense function whereas a small exponent gives a large diffuse function. For electronic wave function calculations, it is convenient to use Slater functions, as they more correctly describe the qualitative features of the MOs than do Gaussian functions, and fewer Slater basis functions than Gaussian basis functions are needed in the basis function expansion of the MOs, for comparable results. Nevertheless, GTOs are universally used in the determination of the two-electron integrals due to

the low computational cost of GTOs against STOs, although a higher number of GTOs is required to have a certain accuracy. Generally, these basis functions are placed at the nuclei or sometimes at the center of a chemical bond.

Classification of Basis Sets

Once chosen the type of the basis functions and the place where they are located in the system, the next main factor to consider is the number of the atomic orbitals to be used. According to this criterion, the basis sets can be classified as:

- **Minimum Basis Set** (MBS) includes only the minimum number of basis functions needed to describe the electrons of a neutral atom. For instance, ($1s$) is the only basis function for the hydrogen atom. For atoms of the first row, ($1s$, $2s$, $2p_x$, $2p_y$, $2p_z$) are the minimum number of basis functions.

- **Double Zeta** (DZ) doubles the number of basis functions. Therefore, ($1s$, $1s'$) are the two s-type functions for the hydrogen atom, and ($1s$, $2s$, $1s'$, $2s'$) and ($2p$, $2p'$) are the four s-type and two p-type functions, respectively, for atoms of the first row.

- **Split Valence Basis** doubles the number of valence orbitals which have an important role in the chemical bond. Thus, in the case of DZ, the basis set is known as VDZ (valence double zeta). Triple, Quadruple, Quintuple Zeta (TZ, QZ, 5Z), and split valence basis (VTZ, VQZ, V5Z) are the following levels of basis sets.

- **Polarization functions** are basis functions with high angular momentum. For instance, d orbitals are used to polarize p orbitals, f orbitals

to polarize d orbitals, and so on. For single determinant wave functions (electronic correlation not included) the first set of polarization basis (p functions for hydrogen atoms and d functions for heavier atoms) is the most important contribution to describe the relevant effects of polarization in the electronic distribution. For basis sets that include electronic correlation, functions with high angular momentum are indispensable to obtain accurate results. When a set of polarization functions is added to a DZ basis set, the new basis set is known as Double Zeta plus Polarization (DZP).

- **Diffuse functions** have small exponents in the basis functions and are used to describe electrons not so much bound (anions, Rydberg states) or to find properties which depend on the tail of the wave function (polarizability).

Balanced Basis Sets

A general rule expresses that the number of basis functions of a given type should be, at least, one less than those whose type corresponds to a previous angular momentum. For instance, $3s2p1d$ is a balanced basis set whereas $3s2p2d2f1g$ is too much polarized. The use of mixed basis sets is not recommendable. Sometimes one can use a DZP basis set for the most important part of the system and a minimum basis set for spectator parts of the system. But some unbalanced results can be obtained instead. In the hydrogen molecule, if one chooses the minimal basis set for one H and a DZP for the other H, the computation will predict a dipole moment in the H_2, since the variational principle will preferentially place the electrons near the centre with the most basis functions. The majority of calculations are therefore performed with basis sets of the same quality on all atoms, possibly cutting polarization and/or diffuse functions on hydrogens. But the same quality of basis functions in the overall system does not guarantee the balanced result. LiF in a DZP basis set for both atoms is not enough to have a good precision in the result, as F has

less basis functions per electron than Li. This effect is minor with increasing number of basis functions. Once the basis set is chosen, the next step is how to select the numerical values of the exponents of these functions. They are determined by a variational HF procedure for s- and p-type functions where the exponents are the variational parameters.

Contracted Basis Sets

One disadvantage in the basis sets obtained by energy optimization is the dependence of the wave function on the core electrons. The $1s$ electrons have an important weight in the total energy and the optimization procedure makes this basis set optimum for core electrons and not for valence electrons. Nevertheless, many properties like polarizability depend on the external region around the nuclei which is less important energetically. As many functions are used to describe core electrons which are energetically important but not chemically is the fundamental base of the contracted basis sets. For instance, 10 s-type functions are optimized for the carbon atom by a variational procedure. It is possible to find that six of them describe the $1s$ orbital whereas the remaining four basis functions describe the $2s$ orbital. The most external region is the most chemically important. As the computational cost is raised to the fourth power of number of the basis functions [99] and the major effort is done for the less important part under chemical point of view, the variational procedure can be performed treating the coefficients in front of the inner basis functions as constants. For the above mentioned example, the contraction of the 10 basis functions can be done as 6,3,1 reducing considerably the number of variational parameters from ten to three. The general expression of the contraction is shown below.

$$\chi(CGTO) = \sum_{i=1}^{k} a_i \chi_i(PGTO) \tag{4.3}$$

Linear combinations of primitive GTOs (PGTOs) with the expansion coefficients fixed give the contracted GTOs (CGTOs). The contraction of basis functions increases the energy so the number of variational parameters is reduced and, therefore, the basis set is less flexible, but it also reduces the computational cost.

An example of primitive and contracted functions for the first row atoms and hydrogen atoms is $(10s4p1d/4s1p) \longrightarrow [3s2p1d/2s1p]$. The basis set within parentheses show the number of primitives and the contracted functions are expressed within square brackets.

There are two manners to contract a set of PGTOs into a set of CGTOs:

- Segmented contraction
- General contraction

The segmented contraction was the first one used in quantum-chemical computations. A set of PGTOs is divided into smaller sets of CGTO basis functions. In the previous example about the contraction of 10 s-type functions into three CGTOs as 6,3,1 can be expressed, according to Eq. 4.3, as:

$$\begin{aligned} \chi_1(CGTO) &= \sum_{i=1}^{6} a_i \chi_i(PGTO) \\ \chi_2(CGTO) &= \sum_{i=7}^{9} a_i \chi_i(PGTO) \\ \chi_3(CGTO) &= \chi_{10}(PGTO) \end{aligned} \tag{4.4}$$

In the segmented contraction each primitive is used in only one contracted function. Nevertheless, in the general contraction all primitives (for a given atom and angular momentum) contribute to all contracted functions with the same angular momentum but different coefficient.

$$\begin{aligned} \chi_1(CGTO) &= \sum_{i=1}^{10} a_i \chi_i(PGTO) \\ \chi_2(CGTO) &= \sum_{i=1}^{10} b_i \chi_i(PGTO) \\ \chi_2(CGTO) &= \sum_{i=1}^{10} c_i \chi_i(PGTO) \end{aligned} \tag{4.5}$$

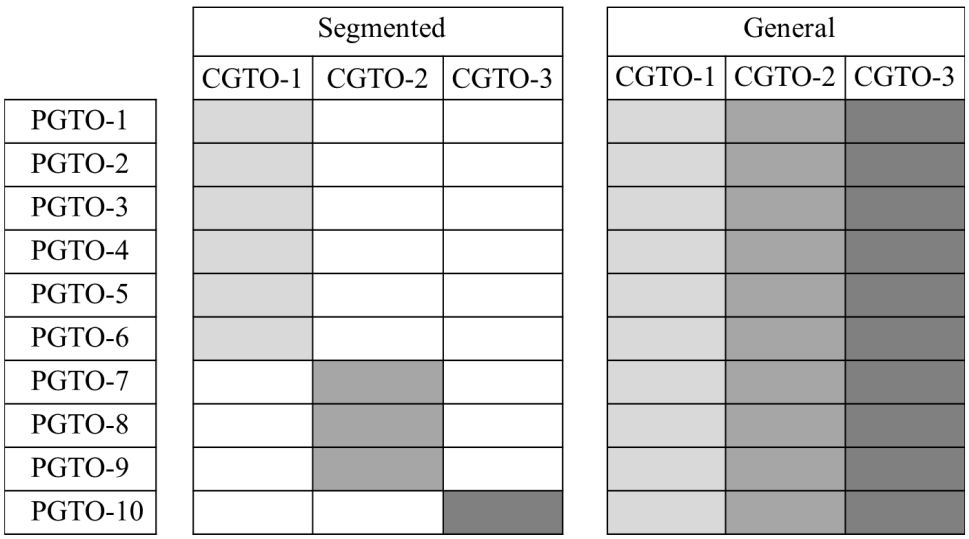


Figure 4.1: Segmented and general contraction scheme for basis sets.

A popular way to obtain the coefficients of the general contraction is by means of Atomic Natural Orbitals (ANO) [128]. The main difference of segmented and general contraction can be shown in Fig. 4.1.

Segmented-contracted Basis Sets

- Pople Basis Sets

- Dunning-Huzinaga Basis Sets

Pople Basis Sets

STO-nG basis sets [129] are Slater type basis sets generated from n PGTOs. The exponents of the PGTOs have been determined fitting them to the STO instead of using variational optimization. n can take values from 2 to 6. STO-3G is the minimum basis set most commonly used. For carbon and hydrogen this basis set is $(6s3p/3s) \longrightarrow [2s1p/1s]$.

k-nmlG basis sets [130] were developed by Pople and co-workers. The value of k shows the number of PGTOs used to describe a core orbital and the label nl represents the number of PGTOs used to double the valence orbitals like in a valence double zeta basis set. When nlm are employed a split-valence triple-zeta basis set is obtained instead. The polarization functions are expressed after G and diffuse functions before G with a plus symbol. The exponents of the PGTOs are determined by variational procedure and are the same for s- and p-type basis functions. This reduces the computational cost but the basis set is less flexible. Some examples of these basis sets are 3-21G, 6-31G, 6-311G, 6-31G(d) or 6-31+G(d). The last one shows that d polarization functions and s- and p-type diffuse basis functions are added to a 6-31G basis set.

Dunning-Huzinaga Basis Sets

Huzinaga developed optimized no contracted basis sets up to $(10s6p)$ for the first row elements [131]. Afterwards, Duijneveldt extended the basis sets up to $(14s9p)$ [132] and Partridge up to $(18s13p)$ [133]. Dunning used the Huzinaga's PGTOs to contract them in different ways giving birth to Dunning-Huzinaga (DH) basis sets [131]. A commonly used DH basis set is $(9s5p/4s) \longrightarrow [3s2p/2s]$ whose contraction scheme is 7,2,1 (where one primitive enters twice) for s-type basis functions, 4,1 for p-type basis functions and 3,1 for hydrogen atoms. The TZ basis set is also broadly used and is obtained from $(10s6p/5s) \longrightarrow [4s3p/3s]$ with the contraction scheme 5,3,1,1 for s-type basis functions, 4,1,1 for p-type functions and 3,1,1 for hydrogen atoms. DH basis sets are more flexible than Pople ones because they do not have the restriction concerned about the same exponents for s- and p-type functions.

General-contracted Basis Sets

The basis functions are the Atomic Natural Orbitals (ANOs) [128] obtained from correlated atomic computations. The natural orbitals resulting from correlated calculations of free atoms at the CISD or MCSCF level diagonalize the density matrix, and their eigenvalues are the orbital occupation numbers. ANO functions describe properly the electron correlation in the atomic problem and should be useful to describe the molecular electron correlation. For each atomic symmetry (s, p, d...), each CGTOs is a linear combination of all PGTOs (general contraction). Different ANO-type basis sets have been contracted from different number of primitives: ANO-S [134] and ANO-L [135, 136, 137]. The basis set ANO-RCC [138] takes into account relativistic effects and is used in heavy atoms. One advantage of ANO basis sets is to be derived from adding the next important atomic natural orbitals regarding their occupation number. This natural criterion to choose basis functions of different symmetries and extend the orbital set according to the contribution to the total one-electron density gives balanced basis sets.

Correlation Consistent Basis Sets

Correlation consistent (cc) basis sets [139, 140] are designed to recover the correlation energy of the valence electrons and are optimized using correlated CISD wave functions. The cc basis sets are referred to as cc-pVXZ which means a Dunning correlation-consistent, polarized valence, X-zeta basis, where X takes labels D, T, Q, 5, 6 for double, triple, quadruple, quintuple, and sextuple functions added in shells. For instance, cc-pVTZ would be $4s3p2d1f$ and cc-pVQZ would be $5s4p3d2f1g$. When 'aug' prefix [141] is used in the basis set, a set of diffuse functions for every angular momentum present in the basis is added. Thus, aug-cc-pVDZ for C atoms has diffuse s-, p-, and d-type basis functions. To describe core and core-valence correlation, tail functions [142] are added to the basis sets which are known as cc-pCVXZ. Thus, for no hydrogen atoms cc-pCVDZ has $1s1p$ tail functions, cc-pCVTZ has $2s2p1d$ tail functions, and so on. Correlation consistent basis are designed to converge smoothly toward the complete basis set limit like ANO-type basis sets. These basis sets have a segmented contraction but some benchmarks have shown that cc-pCVDZ gives 65% of correlation energy, cc-pCVTZ 85%, cc-pCVQZ 93%, cc-pCV5Z 96% and cc-pCV6Z 98%. This systematic tendency allows designing schemes to extrapolate the basis set limit to infinite [143, 144, 145]. The main disadvantage in the complete basis set extrapolation is the high increase of basis functions as the quality of the basis set is better. So, only small systems like ethene can be studied by extrapolated basis sets [146].

Linear dependence problem in the basis functions

The use of a huge number of basis functions, especially when employing diffuse functions, can derive in redundant linear combinations of those basis functions. Then, there is a linear dependence in the basis set which is overcompleted. To evaluate the degree of linear dependence in a basis set, the eigenvalues of the

hermitian **S** overlap matrix, resulting from the Roothaan equations [95, 102], are obtained:

$$\begin{aligned} S_{\alpha\beta} &= \langle \chi_\alpha | \chi_\beta \rangle \\ \mathbf{U^{-1}SU} &= \mathbf{s} \end{aligned} \tag{4.6}$$

where **s** is a diagonal matrix of the eigenvalues of **S**, and **U** is a unitary matrix. If a basis set is linear dependent, at least one eigenvalue is zero and **S** is known as a singular matrix. The smallest eigenvalue of **S** is a good indicator of the linear dependence of the basis set. A singular **S** matrix can derive to numerical problems when solving the SCF equation. Therefore, in practice, there is an upper limit over the completeness of a basis set.

Chapter 5

Photochemical and Photophysical concepts

Spectroscopy is the study of the interaction of electromagnetic radiation with matter. A spectroscopic analysis of the irradiate matter can be useful in many fields of Chemistry [147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157]. Spectroscopic phenomena applied in Chemistry are explained by the Molecular Orbital theory [158]. Considering a number of quantized molecular energy levels in an atomic or molecular system, an electron is promoted from a low energy level to another with higher energy when the interaction of radiation with matter occurs at a certain frequency. If the electron promotes from a high energy level to another with lower energy, the exceeding energy is emitted at a frequency determined by the same relation than in the absorption process. The energy difference between two states is expressed as follows:

$$\Delta E = h\nu \tag{5.1}$$

The lowest energy state of an atom or molecule is the ground state (the most stable). The states of higher energies are known as excited states. A system in the excited state can dissipate the excess energy via three mechanisms: *radiative transition* (emission of radiation), *nonradiative transition* between different states of the molecule in such a way the energy is dissipated as heat or from a molecule to another (quenching), and a *chemical reaction*.

In the simplified molecular orbital model, the transitions between orbitals responsible of the states should be characterized. In fact, molecular orbitals can be classified as core, valence and Rydberg. Valence orbitals can be bonding and antibonding (delocalized over the whole molecule), or lone pairs localized in an atom [149, 158, 159]. A valence excited state is explained as the promotion of an electron from an occupied molecular orbital to a virtual molecular orbital of a molecule. On the other hand, a Rydberg state is obtained when the electron is excited from an occupied molecular orbital to an atomic orbital of increasing n quantum number, whose diffuse nature makes charge to be localized far away the remaining electrons of the molecule. The electron in the Rydberg orbital sees the whole molecule like a cation with a punctual charge. The charge transfer states [160] take place when an electron from an occupied molecular orbital of a molecule is transferred to a virtual molecular orbital of other molecule (or between two different moieties of the same molecule).

According to the energy, the electromagnetic spectrum is divided into different regions:

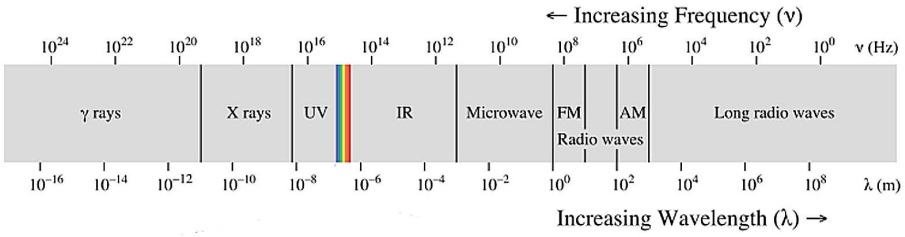


Figure 5.1: Electromagnetic Spectrum

Table 5.1: Electromagnetic spectrum regions of chemical interest and associated molecular phenomena.

Region	**Molecular phenomena**	**Energy** (eV)	**Wavelength** (nm)
X Rays	Excitation of internal electrons	120 - 240×10^3	10 - 6×10^{-3}
Ultraviolet-Visible (UV-VIS)	Redistribution of external electrons (electronic transition)	UV: 3.2 - 120 VIS: 1.7 - 3.2	UV: 390 - 10 VIS: 780 - 390
Infrared (IR)	Change in Configuration (vibrational transiton)	10^{-3} - 1.7	1 - 780
Microwave	Change in Orientation (rotational transition)	4×10^{-5} - 10^{-3}	30cm - 1mm

The interaction of radiation-matter is studied in a semiclassical treatment: matter is considered as a quantized system (with rotational, vibrational and electronic levels), while electromagnetic radiation is treated as a classical wave. Every electronic state has vibrational states close in energy which are associated to interatomic vibrations. And every vibrational state has rotational states even closer in energy resulting from rotations of the molecules around their gravity centers. When radiation interacts with matter, the energy quanta are distributed in the different degrees of freedom or movements of the molecule (translational, rotational, vibrational, electronic, and nuclear, in increasing energy order) [161]. From a photochemical point of view the states of interest are the vibrational and electronic states, so they take place in most of the biological processes.

Before absorbing radiation, molecules are in their ground electronic state. There is a vibrational threshold energy known as zero point energy, which is associated to the minimum-energy state of a molecule. The high-energy vibrational states can be populated thermically. Visible and/or ultraviolet electromagnetic radiation is, however, needed to populate an electronic state different from the ground state.

The absorption of radiation energy from the ground state is used to change the charge distribution of the molecule through the excitation or relocation of their electrons. The molecule is polarized due to the electric field causing a Transition Dipole Moment (TDM) defined as [162]:

$$\mathrm{TDM}_{n\to m} = \sum_{\alpha=x,y,z} \langle \psi_m | \hat{d}_\alpha | \psi_n \rangle \tag{5.2}$$

where $\hat{d}_\alpha$ is the dipole moment operator for component α. The square of the transition dipole moment estimates the probability of the transition $n \to m$, and hence, the band intensity (Fermi's Golden rule). If the dipole moment is zero the one-photon transition is forbidden. The selection rules of allowed one-photon transitions can be deduced by Group Theory. A transition is allowed if and only if the direct product of the irreducible representations of the components, $\Gamma(\psi_m) \otimes \Gamma(\hat{d}_\alpha) \otimes \Gamma(\psi_n)$, is equal to the totally symmetric irreducible representation of the punctual group which the molecule belongs to.

The intensity is a collective phenomenon. When a transition is the most probable to take place, the intensity has its maximum value. The oscillator strength (f) is a measure of the intensity spectral band. It represents the probability of a system to access from the ground state to an excited state by absorbing one photon. For a transition between the initial state n and the final state m, the oscillator strength is directly proportional to the square of

the module of the $\mathrm{TDM}_{n\to m}$ and to the energy difference between both states as shown below:

$$f_{n\to m} = \frac{2}{3}\Delta E_{nm}|\mathrm{TDM}_{n\to m}|^2 \tag{5.3}$$

According to the Franck-Condon (FC) principle, the most likely transition to take place corresponds to that without relative changes in the position of the nuclei of the molecule, since the electronic transition due to the absorption of a photon is very fast as compared to the nuclear movement. This transition is known as *vertical transition*. The excitation of an electron from a quantum state to another higher in energy is the first step of a long path which can lead to a final photochemical product. The electronic states are represented by energy hypersurfaces which constitute the potentials subjected to the molecules due to the different distributions of charge density. These surfaces are known as *Potential Energy Surfaces* (PES).

The excited states populated after the absorption process will have a limited lifetime, since the molecule is unstable regarding the ground state and will search for energy relaxation channels to be deactivated. The energy relaxation can occur through a great number of phenomena which are different to each other in their execution time, i.e. in their probability. The main ways of deexcitation can be classified as radiative or nonradiative processes, as shown in Fig. 5.2.

The absorption is the fastest process, in the femtosecond timescale (10^{-15} s), since the electrons are the only involved in the process and there is no nuclear reorganization. After the absorption between states of the same spin multiplicity (commonly singlet states), most of the molecules are in a vibrational excited state.

The first step of deexcitation is the *Intramolecular Vibrational Relaxation*, IVR, i.e. the energy dissipation occurs through vibrations in the molecule.

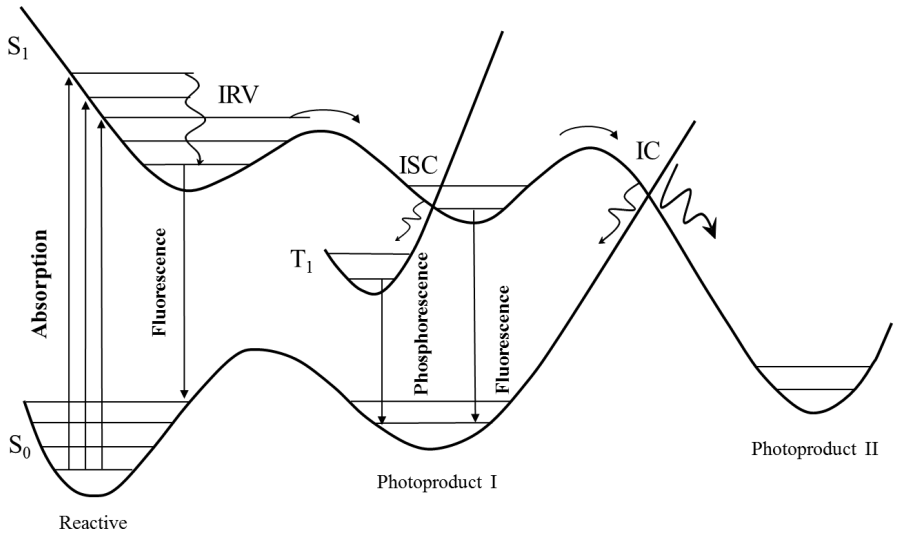


Figure 5.2: Main intramolecular deexcitation mechanisms in Photochemistry.

Other vibrations are characterized by an energy transfer from the molecule to the solvent. This process is known as *Vibrational Energy Transfer*, VET. The vibrational phenomena are very fast ($10^{-14} - 10^{-11}$ s), although the velocity of these processes depends on the molecular size. During this time, the molecule goes through a minimum energy path from the initial structure, in the Franck-Condon region, to the minimum of the excited state. This path can be crossed by potential energy surfaces of the same spin multiplicity (*Internal Conversion*, IC) or with different spin multiplicity (*Intersystem Crossing*, ISC). In the first case, the existence of an effective crossing between both surfaces is accounted for the conical intersection (CI) or avoided crossing (AC) terms. When the energy of the molecule is trapped in the minimum of the excited state, a shorter and less likely process is taken place, which is the radiative emission to other low-energy state. The emission can come from states with equal or different spin multiplicity and it is known as fluorescence or phosphorescence, respectively. The emission can be a multiple phenomenon as shown in Fig. 5.2.

The Kasha's rule states that the emissions occur from the lowest-energy excited state to the ground state, since the rest of relaxation processes are much faster and more likely to happen. The emitted radiation has a longer wavelength than the absorbed by the molecule. This phenomenon is named as *Stoke's shift* [163, 164]. The lifetime of these processes are relatively longer ($10^{-9} - 10^{-6}$ s for fluorescence and $10^{-3} - 10^{2}$ s for phosphorescence) and are dependent on the transition probability between states. Electronic transitions between states of different spin multiplicity, i.e. phosphorescence and intersystem crossing, are forbidden, unless the spin-orbit coupling (SOC) favors the mixture between singlet and triplet states.

Quantum yield (Φ) measures the relation between the absorbed and emitted radiation. Systems with an intense emission have a quantum yield near one. When the quantum yield is decreased, molecules are deactivated by other processes which can be nonradiative (internal conversion or intersystem crossing), radiative (fluorescence or phosphorescence) or quenching. The nonradiative phenomena in a molecule occur when a funnel is found between two potential energy surfaces of the same spin multiplicity (internal conversion) or different spin multiplicity (intersystem crossing), through a hyperline or a hyperplane, respectively. These processes are faster as the states are closer in energy and their lifetimes are similar to those of vibrational relaxation, $10^{-14} - 10^{-11}$ s.

- **Photochemical Reactions**

The global process of deexcitation can give a photochemical reaction, i.e. the ground-state minimum reached by the system after relaxation is different from the initial minimum of the system before the interaction of the molecule with light. Adiabatic and nonadiabatic photochemical processes can occur in the relaxation pathway. Within the Born-Oppenheimer approximation [95, 96, 97, 98], adiabatic photochemistry is characterized by processes which take place in a unique potential energy surface, whereas nonadiabatic photochemical processes start in an excited-state potential energy surface and

finish in a potential energy surface of other electronic state, commonly the ground state. In Fig. 5.3 are shown both types of photochemical reactions. After energy absorption, the reactive R is promoted to an electronic excited state R*. The reaction R* → P in the right-hand side of Fig. 5.3 is an adia-

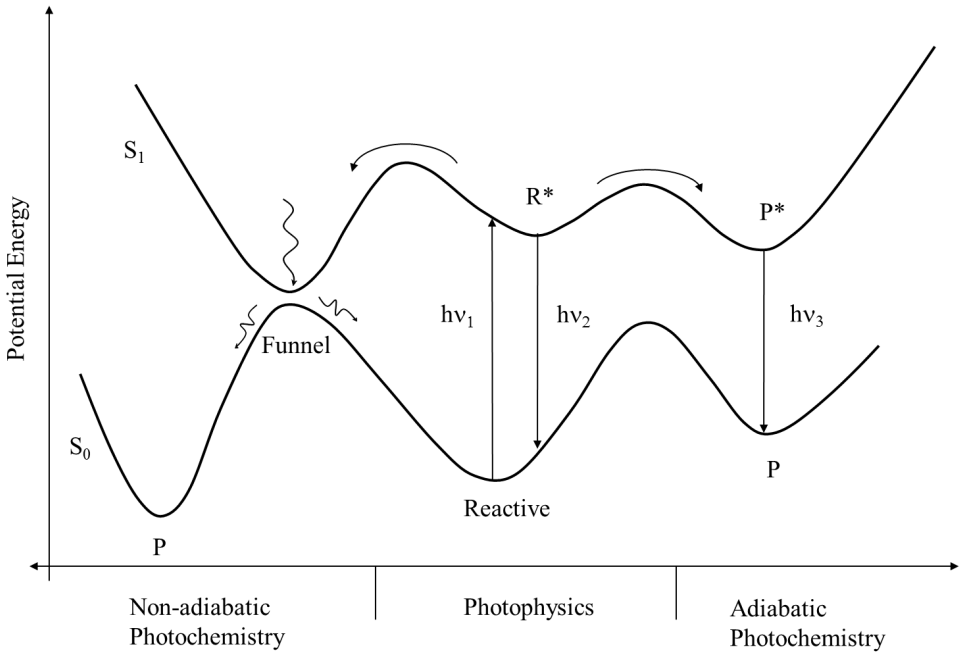


Figure 5.3: Photochemical processes.

batic process leading to an emitting feature whereas the left-hand side is a nonadiabatic process, since the reaction starts in an excited-state potential energy surface followed by a nonradiative decay to the ground-state potential energy surface. The photochemical processes compete with the photophysical phenomena, in the central part of Fig. 5.3, which lead to the regeneration of the starting system. The pathway of the photochemical reaction will be determined by the energy barriers in the excited- and ground-state surfaces. The emissions will occur from the relative minima of the excited-state potential

energy surfaces and the efficiency of a nonradiative process in a crossing of two surfaces is the key point of nonadiabatic photochemistry. In modern photochemistry the efficiency of radiationless decay between different electronic states, taking place in internal conversion and intersystem crossing processes, is usually associated to the presence of conical intersections and crossing hyperplanes which behave as funnels where the probability for nonadiabatic jumps is high [16, 162].

A crossing seam occurs between two states of the same spin multiplicity when their PESs intersect along a $(F-2)$-dimensional hyperline as the energy is plotted against the F nuclear coordinates, where F is the number of internal degrees of freedom $(3N-6)$. In any point of the $(F-2)$-dimensional intersection space the energies of the two states are the same and there is no interaction between both states. The degeneracy is lifted along the two remaining linearly independent coordinates, x_1 and x_2, that span the branching subspace corresponding to the gradient difference vector and the nonadiabatic coupling vector, respectively [16, 162]. See Fig. 5.4.

Gradient difference vector:

$$x_1 = \frac{\partial(E_1 - E_2)}{\partial Q} \tag{5.4}$$

Nonadiabatic coupling vector:

$$x_2 = \langle \psi_1 | \frac{\partial \psi_2}{\partial Q} \rangle \tag{5.5}$$

where Q are the nuclear coordinates of the system.

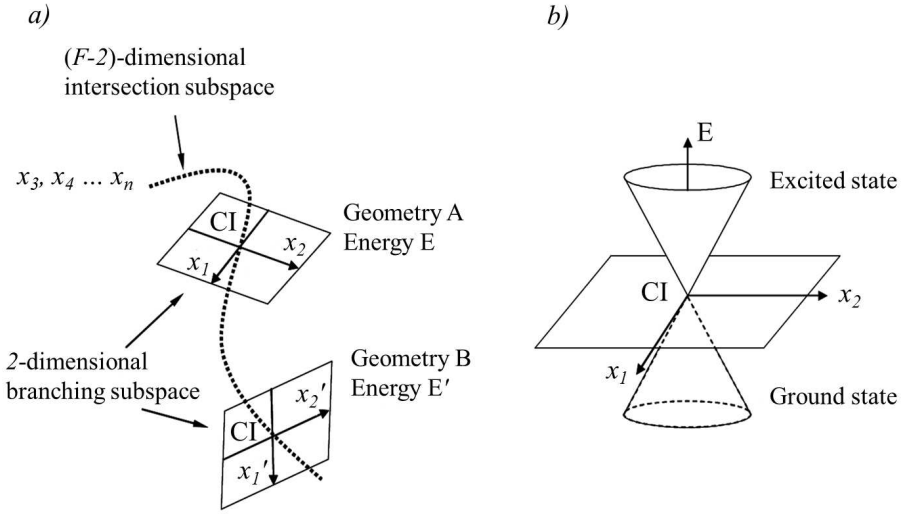


Figure 5.4: Hyperline resulting from the intersection of two PESs (a) and the energy representation of the excited and ground states against x_1 and x_2 coordinates of the branching space (b).

The crossing between states of different spin multiplicity corresponds to a $(F-1)$-dimensional set of points where the intersystem crossing may occur and cause the relaxation of the system to a state of different multiplicity [165]. In this case, the branching space is one-dimensional, since the nonadiabatic coupling vector vanishes, and the intersection space is referred to as an hyperplane.

In order to describe the dynamics of photochemical reactions, full characterization of the hyperline is required. The crossing seam can be viewed as formed by an infinite number of CI points. The denomination conical intersection comes from the fact that the corresponding PESs at a CI point have the shape of a double cone when the energy of the upper and lower states is plotted against the x_1 and x_2 coordinates. Normally, in particular for low-energy processes, transition from one state to the other is expected to take place in the region of the lowest-energy point of the hyperline, i.e. at the lowest CI, which is also

named minimal in the crossing seam (MXS) [166] or minimum energy crossing point (MECP) [167]. Also in reactions involving crossing hyperplanes, one searches for the lowest-energy point along the $(F-1)$-dimensional intersection space [165].

- **Spin-Orbit Coupling**

The spin-orbit coupling (SOC) term in the Hamiltonian accounts for the interaction between the magnetic spin momentum of the electron ($\vec{S}$) and the magnetic momentum due to the electronic orbital movement ($\vec{L}$). The spin selection rules are a consequence of the electric dipole moment operator which does not depend on the spin and, therefore, only transitions with the same multiplicity are allowed. In fact, singlet-triplet absorptions are hardly observable in the spectra of hydrocarbons and other molecules without heavy atoms. Nevertheless, the SOC enables the mixing-up of states with different multiplicity. The expression of the spin-orbit Hamiltonian is shown below [162]:

$$\hat{H}_{SO} = \frac{e^2}{2m_e^2c^2} \sum_j \sum_\mu \frac{Z_\mu}{|\mathbf{r}_j^\mu|^3} \hat{\mathbf{l}}_j^\mu \cdot \hat{\mathbf{s}}_j \tag{5.6}$$

where $\hat{\mathbf{l}}_j^\mu$ is the orbital angular momentum operator of the electron j and $\mathbf{r}_j^\mu$ is the vector pointing from the nucleus μ to the electron j. According to Eq. 5.6, the SOC is relevant for atoms with high atomic number, Z_μ. The $\hat{H}_{SO}$ operator corresponds to a small term in the electronic Hamiltonian, so the singlet states will have a low contribution of triplet states and the same for triplet with a low contribution of singlet states. This makes the singlet-triplet transition momentum different from zero and enhances the probability of the transition $S_i \rightarrow T_f$. The length of the SOC vector is given by the expression:

$$\mathrm{SOC}_{fi} = \sqrt{\sum_{u} |\langle T_{f,u}|\hat{H}_{SO}|S_i\rangle|^2} \qquad u = x, y, z \tag{5.7}$$

Spin-orbit coupling allows to evaluate the efficiency of an intersystem crossing process. Thus, negligible SOC values lead to very low ISC rates [165]. For instance, SOC values of 0.1 cm^{-1} are considered high enough for small organic molecules.

- **Radiative Lifetimes**

In the energy relaxation processes, the kinetic constants (k) and radiative lifetimes (τ_{rad}) determine the dynamics of a photoreaction. The kinetic constant of a decay process is given by the sum of both radiative and nonradiative kinetic constants, as shown in Eq. 5.8, which account for the radiative phenomena (fluorescence or phosphorescence) and the radiationless processes (internal conversion or intersystem crossing), respectively, during the photoreaction.

$$k_{decay} = k_{radiative} + k_{nonradiative} \tag{5.8}$$

The lifetime of an excited state is also defined by both radiative and nonradiative contributions, since its excess energy can be dissipated by different deactivation mechanisms. Similarly to Eq. 5.8, the reciprocal of the lifetime is given by the expression below:

$$\frac{1}{\tau_i} = \frac{1}{\tau_{radiative}} + \frac{1}{\tau_{nonradiative}} \tag{5.9}$$

where the i index labels singlet or triplet state. As aforementioned, the Kasha's rule states that the fluorescence process occurs from the lowest-energy singlet

excited state (S_1) to the singlet ground state (S_0) whereas the mechanism of phosphorescence is provided by the spin-orbit coupling between the lowest triplet state (T_1) and the singlet ground state. Experimentally, the measured lifetimes of radiative features are the same as the corresponding excited-state lifetimes, i.e. fluorescence for the S_1 state and phosphorescence for the T_1 state [168]. On the other hand, the radiative lifetime (τ_{rad}) is defined as the reciprocal of the kinetic constant associated to the radiative channel (k_{rad}), without considering the nonradiative phenomena involved in the decay process. Thus, the fluorescence and phosphorescence quantum yields (Φ_j) can be rewritten in terms of the excited-state (τ_j) and radiative ($\tau_{\mathrm{rad},j}$) lifetimes, where the j index refers to the type of radiation.

$$\Phi_j = \frac{k_{\mathrm{rad},j}}{k_j} = \frac{\tau_j}{\tau_{\mathrm{rad},j}} \tag{5.10}$$

For the radiative emission, the rate constant k_{rad} both in fluorescence and phosphorescence is related to the Einstein coefficient for spontaneous emission, A_{mn}, and the use of the Strickler-Berg relationship [169] allows to find theoretically the radiative lifetime [170].

$$k_{\mathrm{rad}} = A_{mn} = \frac{1}{\tau_{\mathrm{rad}}} = 2.142005 \times 10^{10} \left(\mathrm{TDM}_{n\to m}\right)^2 \left(\mathrm{E_{VE}}\right)^3 \tag{5.11}$$

where the transition dipole moment is found for the one-photon allowed absorption transition $n \to m$ and $\mathrm{E_{VE}}$ is the vertical emission energy, although the electronic band origin energy ($\mathrm{T_e}$) is used instead of $\mathrm{E_{VE}}$ for being more representative of the emission energy. Strickler-Berg relationship is only valid under the hypothesis that excited-state nonradiative deactivation is of minor importance. For transitions of different spin multiplicity, the l-component of the transition dipole moment depends on the SOC between the triplet and ground states according to the equation:

$$\mathrm{TDM}_l(T^k) = \langle S|\hat{r}^l|T^k\rangle = \sum_n \frac{\langle S^0|\hat{r}^l|S_n^0\rangle\langle S_n^0|\hat{H}_{SO}^k|T^{k,0}\rangle}{E(T^0) - E(S_n^0)} + \\ + \sum_m \frac{\langle S^0|\hat{H}_{SO}^k|T_m^{k,0}\rangle\langle T_m^{k,0}|\hat{r}^l|T^{k,0}\rangle}{E(S^0) - E(T_m^0)} \quad (5.12)$$

where k labels the magnetic spin sublevels, $l \in (x, y, z)$, $T^{k,0}$ and T^k denote the zeroth and first order wave functions of perturbation theory, respectively, and $\hat{H}_{SO}^k$ is the kth component of the SOC operator [171].

- **Electronic Energy Transfer**

Electronic Energy Transfer (EET) is a process by which an excited-state molecule (donor) transfers its energy to a neighboring ground-state molecule (acceptor). There are two types of electronic interaction to explain an EET process: electron exchange interaction [172] and Coulomb or dipole-dipole interaction [173]. The first mechanism is explained as a double electron substitution reaction, i.e. the electron initially on the donor D* jumps to the acceptor A simultaneously with the jump of an electron on A to D* producing the transition D*A $\rightarrow$ DA*. The exchange resonance interaction occurs via overlap of electron clouds and requires physical contact between the interaction paterns. On the other hand, the Coulomb interaction represents an action at a distance, i.e. the electrons initially on D* stay on D while the electrons initially on A stay on A*. The same transition D*A $\rightarrow$ DA* occurs but in this case the Coulomb resonance takes place via the electromagnetic field and does not require physical contact of the interacting paterns. The classical view of this mechanism involves the induction of a dipole oscillation in A by D* [168].

The trivial or radiative mechanism is other way to explain an EET process. This mechanism consists in the emission of a quantum of light by one molecule

which is followed by the absorption of the emitted photon by a second molecule. The latter species does not at all influence the emission ability of the first molecule, but merely intercepts the emitted photon before being observed. The trivial mechanism is a two-step sequence given below

$$\begin{aligned} \mathrm{D}^* &\rightarrow \mathrm{D} + h\nu \\ h\nu + \mathrm{A} &\rightarrow \mathrm{A}^* \end{aligned}$$

Singlet-singlet energy transfer is a spin-allowed process for both the Coulomb and exchange interactions.

$$\mathrm{D}^*(S_1) + \mathrm{A}(S_0) \rightarrow \mathrm{D}(S_0) + \mathrm{A}^*(S_1)$$

Energy transfer from the donor (singlet) to the acceptor (singlet) can occur either by a long-range Coulomb (dipole-dipole) mechanism or by a short-range electron exchange mechanism. Depending on the distance between the donor and acceptor, the mechanism selected for the singlet-singlet energy transfer follows the scheme of Fig. 5.5.

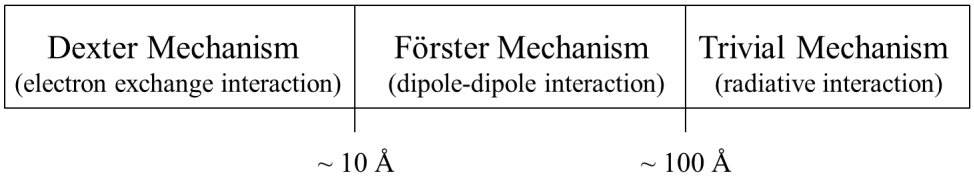


Figure 5.5: EET mechanisms preferred according to the distance between the donor and acceptor.

Triplet-triplet energy transfer is forbidden by the dipole-dipole mechanism but spin-allowed by the exchange mechanism.

$$\mathrm{D}^*(T_1) + \mathrm{A}(S_0) \rightarrow \mathrm{D}(S_0) + \mathrm{A}^*(T_1)$$

An electronically excited donor in its triplet state produces an electronically excited acceptor in its triplet state. In the exchange mechanism, the electronic coupling is not the only key factor that determines the efficiency of the EET processes, but also the resonance condition, i.e. the energy available in the donor must be at least equal or higher than that required to populate the excited state of the acceptor. If this is the case, the process is usually controlled by diffusion and described as exothermic. In the opposite situation, i.e. if the energy of the acceptor is lower than that of the donor, the process becomes thermally activated and lies in the endothermic region [174].

Triplet-triplet energy transfer (TET) is the most common and most important type of energy transfer involved in organic photochemistry [168]. The longer a molecule remains in an excited state the greater the probability that it will transfer electronic excitation energy to a suitable neighbor of its vicinity. In general, the lowest triplet state of a molecule is longer-lived than the corresponding lowest excited singlet state of the same molecule. Therefore, the triplet state of a molecule is a more likely candidate to participate in energy-transfer processes than the singlet state.

Thus, TET is a process where two molecules or molecular fragments interchange both electronic energy and spin. When a closed-shell system is excited, a singlet state will be populated. A triplet state can be accessible via ISC originating the initial triplet state of the donor required in TET processes. This happens, for instance, in the $B_{18}H_{22}$ isomers, studied in the present thesis, which have a non-zero quantum yield of singlet oxygen O_2 ($^1\Delta_g$) production. Schematically, the process is expressed as

$$Borane \xrightarrow{h\nu} {}^1Borane^* \xrightarrow{\text{ISC}} {}^3Borane^*$$

From ${}^3Borane^*$ the singlet oxygen can be produced in the presence of molecular oxygen ${}^3O_2$ ($^3\Sigma_g^-$) through a triplet-triplet energy transfer.

$$^3Borane^* + {}^3O_2 \xrightarrow{\text{TET}} Borane + {}^1O_2^*$$

According to spin conservation rules, the resulting species can be one molecule in a singlet state and the other one in a singlet, a triplet or a quintet state [162]. In the present case, both molecules are in a singlet state. The species $^1O_2^*$ is highly reactive and can be used to destroy cancerous cells. This is the basic principle of photodynamic therapy (PDT) [175, 176].

- **Technical and computational features**

The variational principle ensures that the use of a more extended basis set in the energy computation of a molecular system gives an energy closer to the exact value [95]. Moreover, extended basis sets define a flexible wave function resulting in a better description of the molecule. Unfortunately, the increase in the size of the basis sets and the use of methods which treat properly the electronic correlation can lead to a rising computational cost. Therefore, for instance, it is common to use low-level methods to localize geometries of singular points (minima, transition states, conical intersections) and, then, improve the energy with the employment of high-level methodologies. Differential correlation effects can, however, appear in this procedure. In order to show this fact, the steps followed in the computation of conical intersections are explained in detail. Determination of CIs are usually carried out by using variational wave functions, primarily of the complete active space self-consistent field (CASSCF) and the multireference configuration interaction (MRCI) type, through two main algorithms based on Lagrange multipliers [166, 167, 177, 178, 179, 180] and on projected gradient techniques [181, 182]. A typical computational strategy is to perform single-point multiconfigurational second-order perturbation theory (CASPT2) [9, 109, 183] calculations at CASSCF determined conical intersection geometries, that is, the protocol CASPT2//CASSCF. This procedure is only valid when the hypersurfaces described at both levels of theory behave more or less parallel with a constant relative influence of dynamical correlation [96, 164]. In many situations, because

of differential dynamical electron correlation effects which are not accounted for at the CASSCF level, the CASPT2//CASSCF protocol leads to unphysical results. In Fig. 5.6 the differential correlation effects in the localization of a conical intersection by using the CASPT2//CASSCF protocol are shown. When the PESs computed at the CASSCF and CASPT2//CASSCF level show a parallel behavior (case a), the geometries computed at the CASSCF level are considered to be correct and, therefore, the dynamical correlation contributes in a regular way to both states. In such situations, the CASPT2//CASSCF protocol can be applied confidently. Nevertheless, if the dynamical correlation is different for both states and vary significantly along the PES, geometry optimizations must be performed at the highest level of theory (case b).

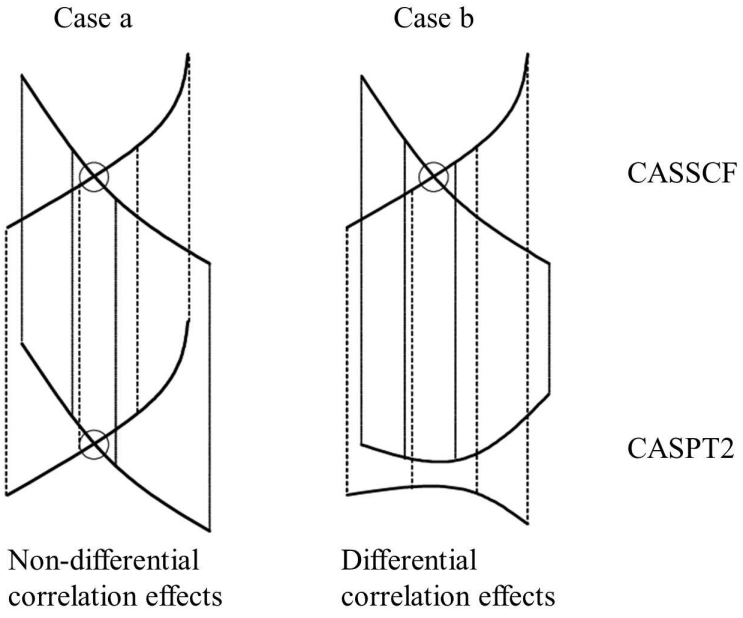


Figure 5.6: Correlation effects in the CASPT2//CASSCF protocol.

As multiconfigurational perturbation theory has succeeded to describe accurately excited states, CASPT2 method can be used to compute CIs. One of the main problems of using perturbation theory to compute PES crossings is that the resulting wave functions corrected up to first order are not, in general,

orthogonal. This drawback can be easily overcome by using MS-CASPT2 method. But this procedure is only valid when the off-diagonal coupling terms are similar and small. This condition can be achieved by enlarging the active space, otherwise artificial large energy splittings are yield [16]. Furthermore, multi-state treatment is used in the computation of Rydberg states, and the level shift technique [12], available in the perturbative method, is employed in order to avoid weakly coupling intruder states interference. The IPEA correction of the zeroth order Hamiltonian is taken into account in diverse studies of the present thesis [14].

Programs

There are many programs that have implemented several quantum, classical and hybrid methods introduced in chapter 3 in order to determine theoretically the energy and properties of molecular systems. The software package most used in the present thesis is MOLCAS [7], a modular program mainly designed to the study of excited states. Cholesky decomposition is employed in the SEWARD module of this program to compute the two-electron integrals [184] and the Douglas-Kroll-Hess Hamiltonian is available in this module to include the scalar relativistic effects for heavy atoms [185, 186]. RASPT2 and CASPT2 energy calculations, transition dipole moments (TDM), and CASSCF geometry optimizations have been computed with MOLCAS. From the CASSCF transition dipole moments and CASPT2 excitation energies, oscillator strengths (f) and radiative lifetimes (τ_{rad}) have been obtained, the latter using the Strickler-Berg relationship [169, 170]. Spin-orbit coupling terms have been calculated within the AMFI framework [187] and states interaction CAS method (CASSI) [188], included in the MOLCAS package. Recently, the ESPF module has been implemented in MOLCAS to perform, together with TINKER program [189], studies with hybrid methodology. To generate the DNA sequence for QM/MM calculations, the X3DNA program [190] has been employed. The dynamical calculations have been carried out with NEWTON-X program [124] in combination with DYNAMIX module implemented in MOLCAS. Additionally, GAUSSIAN package [191] has been

used for ground-state and triplet-state geometry optimizations computed at the DFT/B3LYP level of theory. The visual analysis of both the computed structures (e.g., minima, minimum energy crossing points, etc.) and molecular orbitals have been performed with MOLDEN [192], MOLCAS-GV and GABEDIT [193].

Chapter 6

Results

6.1. The RASPT2 method for Electronic Excited States

The complete active space to second-order perturbation theory (CASPT2) was developed in the beginning of the 1990s by Roos and co-workers [9] to study those electronic structure cases which required a multiconfigurational description of the reference wave function such as bond breakings and dissociations, potential energy hypersurface degeneracies, symmetry breaking problems, biradical situations, etc. In particular, this method showed to be best suited to deal with the quantum chemistry of the excited state [10]. The inclusion of the correlation effects in the multiconfigurational CASPT2 approach opened the door for studying spectroscopic and photochemical phenomena in systems in which computationally costly approaches such as multireference configuration interaction (MRCI) could not be applied. Nevertheless, some drawbacks in the method are mainly related to the limitation in the size of the complete active space (CAS) due to the difficulties in handling large high-order density

matrices. The maximum active space size of 13-15 molecular orbitals (MOs) is rapidly reached, especially when a large number of Rydberg orbitals are required or the need to include in the CAS a second d-shell ($4d$) for first-row transition metal atoms.

The restricted active space to second-order perturbation (RASPT2) theory [2, 3, 6, 7] emerged to extend the applicability of the CASPT2 method to treat larger systems where the active spaces require to be enhanced. As mentioned in section 1.1, in the CAS framework [1], the MO space is divided into three subspaces with a varying number of electrons: inactive (always doubly occupied), active (with varying occupation from zero to two), and secondary (always empty). All possible excitation levels compatible with spatial and spin symmetry involving the electrons in the active space form the multiconfigurational CAS-CI space used as a reference for a further perturbative CASPT2 treatment [9, 13]. The configurational space rapidly grows to many millions of determinants with the size of the active space, making the treatment unaffordable. The RAS method [2, 3] further divides the active space into three subspaces: RAS1, RAS2, and RAS3. The final multireference space is built by allowing in RAS2 the same type of full CI expansion as previously in CAS, but restricting in RAS1 and RAS3 the excitation level to a predefined range: up to single (S), double (SD), triple (SDT), quadruple (SDTQ), etc., by limiting the number of allowed holes (RAS1) and particles (RAS3). The number of possible divisions of the active space combined with the different allowed levels of excitation increases the choices of configurational expansions and the number of solutions, thus making the RASSCF/RASPT2 method less systematic than the CASSCF/CASPT2 method.

The selection of RAS spaces requires very careful calibration. Finding reliable strategies for general purpose calculation is highly recommended. The RASPT2 method has been tested in the determination of the singlet-triplet state energy splitting of three copper-dioxygen and two copper-oxo complexes [6, 194], and one-electron ionization potential and optical band gaps of ethylene, acetylene, and phenylene oligomers [195]. The approach has been also used to compute potential energy hypersurfaces [24, 196], dissotiation energies,

vibrational frequencies, and bond distances [196]. The RASSCF/RASPT2 method has been applied to obtain the electronic excitation energies of first-row transition metal systems [197], first-row transition-metal complexes of corroles [198, 199], in actinide chemistry [200], Ru(II)bipyridil complex [201], and bonding of one or two NO molecules to Cu(I) in a zeolite environment [202]. It has been shown how RASPT2 offers a similar accuracy, when compared to CASPT2, at significantly reduced computational expense, whereas more demanding calculations out of reach for CASPT2 can be performed with the new formulation.

In this section, the RASPT2 method is tested in the computation of the electronic excitation energies for a set of molecules in order to check the accuracy of the method and the computational strategies in systems and problems of various classes, including valence and Rydberg, singlet and triplet, ligand-field, and charge-transfer excited states in different organic and inorganic systems. Fig. 6.1 shows the set of molecules considered in the study.

The section is divided into three subsections each of which focuses on one aspect of the calibration. Initially, free base porphin will be used as an example of the use of RASPT2 for a molecule with an extended π system whose inclusion in the active space is out of reach for CASPT2. Then, excitation energies of singlet and triplet, valence and Rydberg states of ethene and benzene will be computed in order to test the accuracy and ability of RASPT2 to deal with the simultaneous calculation of valence and Rydberg states, and how it takes care of the valence-Rydberg problem. Finally, the nickel atom and the copper tetrachloride dianion will be computed in order to establish the accuracy and proper strategies required in RASPT2 to handle the required inclusion of a second correlating d shell for first-row transition metal compounds, and how the new method simplifies the calculations and extends their possibilities.

In the present study, CASSCF/CASPT2 and RASSCF/RASPT2 calculations have been carried out for the set of molecules displayed in Fig. 6.1. For ethene and benzene the calculations have been performed with an ANO-L basis set

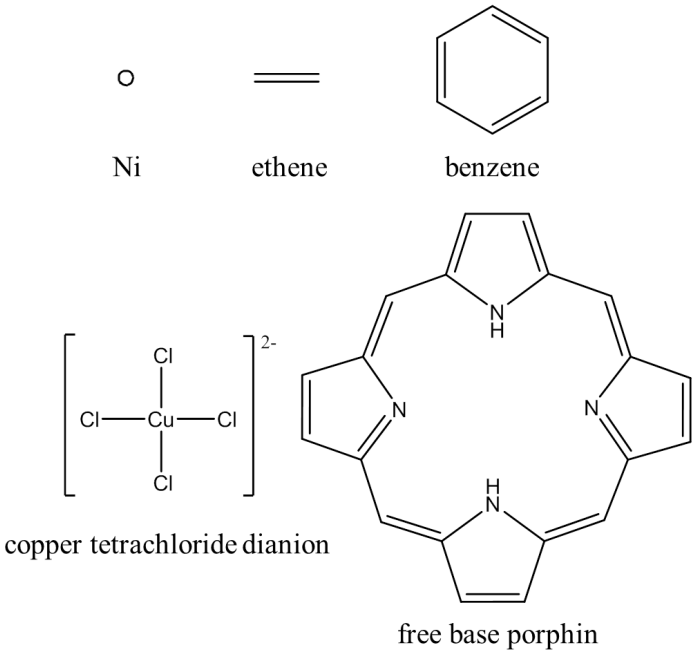


Figure 6.1: Set of molecules considered in the study.

[203] contracted to [4s3p1d] for carbon and [2s1p] for hydrogen atoms. In addition, 1s-type, 1p-type, and 1d-type contracted functions, with diffuse coefficients described elsewhere [15], have been added to this basis set and placed in the centre of the molecule to describe Rydberg orbitals. For free base porphin an ANO-S C,N[3s2p1d]/H[2s] basis set has been used. Cholesky decomposition of the two-electron integrals has been accomplished with a threshold of 10^{-5} au [184]. For the calculations on transition metal systems, ANO-RCC basis sets [204] have been used, contracted to [7s6p4d3f2g] for nickel, [7s6p4d3f2g1h] for copper, and [5s4p2d1f] for chlorine atoms. Scalar relativistic effects have been included using a Douglas-Kroll-Hess Hamiltonian [185, 186]. In ethene and benzene the ground-state geometries have been taken from gas-phase experimental determinations [205]. $CuCl_4^{2-}$ is square-planar (D_{4h}) and the Cu-Cl distance, 2.291 Å, has been taken from a previous study

[206]. At the optimized geometries subsequent single point CASSCF/MS-CASPT2 and RASSCF/MS-RASPT2 calculations have been performed using the mentioned basis sets. When high symmetry is required, for instance D_{6h} in benzene, D_{4h} for $CuCl_4^{2-}$ or spherical symmetry in the nickel atom, the calculations have been performed in a lower-symmetry point group and MOLCAS tools have been used to obtain the proper orbital symmetry. An imaginary level shift [12] of 0.1 au has been used to prevent for weakly coupling intruder states interference and the default shift for the IPEA zeroth-order Hamiltonian [14] (0.25 au) has been employed except for the free base porphin calculations where the IPEA shift has been set to zero to compare the present results to those obtained prior the introduction of the IPEA in 2004 [207, 208]. In all calculations, the core electrons have been kept frozen in the perturbative calculations, except for the nickel atom where the $3s$, $3p$ electrons have been included.

It is important to describe the notation employed to label CAS and RAS calculations. In the first case, the traditional label is used, that is CAS(n, i) where n is the number of electrons included in the active space and i the number of active orbitals. For RAS calculations, a longer notation is used, RAS$(n, l, m; i, j, k)$, where n is the number of active electrons, l the maximum number of holes allowed in RAS1, and m the maximum number of electrons to enter in RAS3. Active orbitals are labeled by i, j, k and refer to those placed in RAS1, RAS2, and RAS3, respectively. Sometimes, the notation S, SD, SDT or SDTQ will be used to emphasize the maximum RAS1 $\rightarrow$ RAS3 excitation level. The active space employed for the various systems will be described in each subsection.

6.1.1. Free Base Porphin

Free base porphin (FBP) is an example of an extended π-conjugated system having 26 valence π electrons and 24 $\pi\pi^*$ MOs (26/24). Including a full valence $\pi\pi^*$ active space is out of reach for a conventional CASPT2 calculation.

Previous studies were performed at the CASPT2(4,4) and CASPT2(16,14) levels [207, 208]. In the former case only four singlet states were computed, whereas eight singlet and eight triplet states were obtained at the latter level of theory. An overall agreement of 0.2-0.3 eV with respect to experimental values was obtained for all eight singlet states belonging to the porphin Q, B, N, and L bands, although in all these cases CASPT2 yields too low values. FBP will be employed as a typical example of how to properly select the RAS active spaces and establish a RAS1/RAS3 excitation level yielding balanced and accurate excitation energies.

Table 6.1 displays a comparison between the new RASPT2 calculations and the previous CASPT2 calculations. The former includes all $\pi\pi^*$ valence electrons and MOs (26/24) in the RAS active space. As in many other organic molecules, the two highest-lying occupied MOs (HOMO and HOMO-1) and the two lowest-lying MOs (LUMO and LUMO+1), are the four most relevant MOs to describe the four lowest-lying states of the molecule [10]. This active space (named Gouterman's space in FBP) was previously used for CASPT2 calculations and showed to be necessary to describe the nature of such states. Indeed, CASPT2(4/4) calculations provided reasonably accurate values for the mentioned states, as well as CASPT2(16/14), including Gouterman's MOs plus other additional orbitals which allowed calculation of higher roots.

In RAS calculations, the most important point is how to partition the RAS spaces in order to include FBP full $\pi\pi^*$ space and obtain accurate results. Not all partitions are equally adequate, and especially the choice of RAS2 has to be made carefully. Table 6.2 summarizes the natural orbital occupation numbers for a number of relevant MOs obtained in a RASSCF(26,2,2;11,4,9)(SD) calculation. This level of theory, including the four Gouterman's MOs in RAS2, the remaining valence $\pi\pi^*$ occupied and unoccupied MOs in RAS1 and RAS3, respectively, and up to double excitations (SD) for the latter spaces, is not intended to get accurate results for all nine computed states, but just to give a guidance in designing the RAS partition. For each excited state the most relevant MOs (see last column in Table 6.2), namely those in which the occupation number is below 1.9 or above 0.1, has been selected carefully.

Table 6.1: Excitation energies (eV) of the singlet and triplet valence $\pi\pi^*$ states of Free Base Porphin (D_{2h}).

State	CASPT2 (4/4)[a]	CASPT2 (16/14)[b]	RASPT2 (26,2,2;11,4,9) (SD)[c]	RASPT2 (26,3,3;11,4,9) (SDT)[c]	RASPT2 (26,2,2;i,j,k) (SD)[e]	STEOM-CCSD[f]	Exp[g]
1^1B_{3u}	1.70	1.63	2.18	1.91	2.18	1.75	1.98 - 2.02 (Q_x)
1^1B_{2u}	2.26	2.11	2.38	2.16	2.38	2.40	2.33 – 2.42 (Q_y)
2^1B_{2u}	2.91	3.08	3.23	2.86	3.23	3.62	3.13 – 3.33 (B)
2^1B_{3u}	3.04	3.12	3.21	3.16	3.21	3.47	3.13 – 3.33 (B)
3^1B_{2u}	–	3.42	5.22 (3.30)[d]	3.37	3.80	4.35	3.65 (N)
3^1B_{3u}	–	3.53	5.38 (3.21)[d]	3.28	3.48	4.06	3.65 (N)
4^1B_{2u}	–	3.96	5.95 (4.02)[d]	4.10	4.20	5.00	4.25 (L)
4^1B_{3u}	–	4.04	6.04 (4.14)[d]	4.22	4.23	5.17	4.25 (L)
1^3B_{2u}	–	1.52	1.83	1.70	1.83	1.26	1.58
1^3B_{3u}	–	1.85	1.99	1.77	1.99	1.80	
2^3B_{3u}	–	1.88	1.98	1.88	1.98	1.98	
2^3B_{2u}	–	1.98	1.98	1.90	1.98	1.85	
CSF[h]	8	537705	63258 (1877432)[d]	1877565	279974/676297	–	–

[a]CASPT2(4,4)/ANO-L $3s2p/2s$, Ref. [207]. Gouterman's four-electron/four-MO CAS space. [b]CASPT2(16,14)/ANO-S $3s2p1d/2s$, Ref. [208]. [c]Present RASPT2 results. Full $\pi\pi^*$ 26-electron/24-MO RAS employed. Gouterman's 4/4 space placed in RAS2. SD or SDT for all states except when indicated. The poor results for the highest singlet states explained in the text. [d]Present RASPT2 results. Full $\pi\pi^*$ 26-electron/24-MO RAS employed. Gouterman's 4/4 space placed in RAS2. Within parentheses are results using SD for the ground 1^1A_g state and SDT for the excited state and CSFs for the excited-state $^1B_{2u}$ SDT calculations. [e]Different RAS spaces partition following the occupation number criterion of Table 6.2. See text. [f]STEOM-CCSD/SVZP results from Ref. [209]. [g]See data in Ref. [208]. [h]Number of Configuration State Functions (CSF) for the 1^1A_g symmetry. In the sixth column CSFs for active spaces (26,2,2;8,6,9)(SD)/(26,2,2;6,8,9)(SD) as examples.

Obviously such number may vary at the different RASSCF levels, but just slightly. It is shown, for instance, that for the four lowest-lying states just the four Gouterman's MOs fulfill such requirements, as expected, whereas two more occupied MOs are required for the 3^1B_{2u} and 3^1B_{3u} states, and two and three more for the 4^1B_{2u} and 4^1B_{3u} states, respectively.

This analysis is important to have a balanced and accurate energy difference between states, the MOs strongly differing in occupation number for such states must be placed in RAS2. That is, to get balanced RASPT2 excitation energies from the ground to the 1^1B_{2u}, 2^1B_{2u}, 1^1B_{3u}, and 2^1B_{3u} excited states, at least the four Gouterman's MOs ($b_{1u}b_{2g}b_{3g}a_u/e^-$;1111/4) must be placed in RAS2. Otherwise the description of the various states will be strongly unbalanced at the initial RASSCF level, and RASPT2 may not be able to recover the desired accuracy. This is better seen in the case of the higher-lying states. In Table 6.1 excitation energies at the RASPT2 (26,2,2;11,4,9)(SD) level of calculation are reported. Only the four Gouterman's MOs are included in RAS2, and single and double excitations (SD) are allowed from RAS1 to RAS3 to obtain the RAS-CI expansion. This level is clearly adequate to describe the four lowest-lying states, largely reducing the computational cost with respect to CASPT2(16/14) (only 10% of CSFs required for the RAS calculations). The case is quite different for the four next states, which require additional MOs to be properly described (see Table 6.2). The RASPT2 excitation energies deviate toward high values by more than 1.5 eV, showing the underestimation of the correlation energy for the excited states as compared with the ground state. Within parentheses, the results of increasing the CI excitation level are included only for the excited states to SDT, while keeping SD for the ground state, a strategy that partially restores the lost balance, giving excitation energies within 0.2-0.3 eV from the experimental values. Similar results are obtained if the level of excitation in RAS1/RAS3 is increased to triple excitations for all states with the RASPT2 (26,3,3;11,4,9)(SDT) calculations, proving that the ground state treatment does not improve with respect to the SD level. In any case, the computational cost increases enormously by including the triple excitations (30 times more CSFs are required).

Table 6.2: Natural occupation numbers of the most relevant molecular orbitals of the low-lying $\pi\pi^*$ states of Free Base Porphin (D_{2h})[a].

State	$3b_{1u}$	$4b_{1u}$	$2b_{2g}$	$3b_{2g}$	$3b_{3g}$	$5b_{1u}$[b]	$2a_u$[b]	$4b_{2g}$[b]	$4b_{3g}$[b]	$3a_u$	RAS2[c] $b_{1u}b_{2g}b_{3g}a_u/e^-$
1^1A_g	1.9692	1.9596	1.9714	1.9561	1.9581	1.8539	1.8631	0.1481	0.1595	0.0543	
1^1B_{3u}	1.9699	1.9606	1.9349	1.9715	1.9646	1.4726	1.4449	0.5448	0.5664	0.0739	1111/4
1^1B_{2u}	1.9696	1.9622	1.9702	1.9601	1.9625	1.5462	1.3962	0.6117	0.4803	0.0626	1111/4
2^1B_{2u}	1.9609	1.9684	1.9662	1.9582	1.9638	1.3169	1.4744	0.5389	0.6988	0.0615	1111/4
2^1B_{3u}	1.9713	1.9527	1.9526	1.9722	1.9451	1.3941	1.4197	0.6275	0.6037	0.0602	1111/4
3^1B_{2u}	1.9906	1.2519	1.9907	1.9637	1.7919	1.7860	1.7749	0.2467	1.0048	0.0992	2121/8
3^1B_{3u}	1.9922	1.3057	1.9617	1.9923	1.7339	1.8496	1.6622	1.1639	0.1479	0.0773	2121/8
4^1B_{2u}	1.2515	1.9914	1.7929	1.9747	1.9898	1.8367	1.7494	0.1429	1.0928	0.0765	3311/10
4^1B_{3u}	1.3231	1.9910	1.9694	1.7753	1.9894	1.8312	1.6278	0.9479	0.2936	0.1310	3312/10

[a]RASSCF(26,2,2;11,4,9)(SD) level of calculation. [b]Orbitals of the 4/4 Gouterman's space. [c]Orbitals and electrons within RAS2 selected from the occupation numbers (<1.9 and >0.1).

More elaborated calculations have been performed for each pair of states (here, the ground and each excited state) which has been computed using the specific active space suggested by the occupation numbers in Table 6.2. In this procedure, both the ground and excited states have in RAS2 those MOs largely changing their occupation number in the excitation process. These calculations (which are equivalent to the RASPT2 (26,2,2;11,4,9)(SD) results for the four lowest-lying states) provide the most accurate set of results for the different states at an intermediate computational cost.

The main conclusion obtained from these sets of calculations on FPB is that RASPT2 can provide accurate results for excited states only if the design of the RAS partition, and particularly the composition of the RAS2 space, is carefully controlled. RAS2 must contain those MOs that largely change their occupation number in the states under comparison. Otherwise, the corresponding states will have an unbalanced CI description, and perturbation theory might be unable to provide accurate excitation energies. Any initial RASSCF SD calculation on the requested states including a large enough active space will be sufficient to identify the MOs that should be placed in RAS2, and the occupation number criterion (<1.9 and >0.1) can be used for guidance. If the RAS2 partition is correct, the single and double (SD) level of CI excitation required in RAS1 and RAS3 is sufficient to provide accurate excitation energies at a reasonable computational cost. Increasing the excitation level (triple or quadruple CIs) may partially compensate the lack of balance, but it typically gives large CI expansions that may become very expensive. The use of the full $\pi\pi^*$ 26/24 active space increases the accuracy compared to more limited active spaces. Furthermore, RASPT2 compares well with experiment, unlike CCSD, especially for the higher states, which deviate from experiment almost 0.8 eV at the CCSD level of calculation. RASSCF/RASPT2 can therefore be considered a very convenient tool to study the spectrum of this type of π-extended systems. This is even more important when carrying out geometry optimizations.

6.1.2. Ethene, Benzene and the Valence-Rydberg mixing problem

Ethene is usually described by two $\pi\pi^*$ valence orbitals - the HOMO (Highest Occupied Molecular Orbital) and LUMO (Lowest Unoccupied Molecular Orbitals) MOs - that form the basis for the low-lying valence singlet and triplet $\pi\pi^*$ states. Additionally, series of diffuse states of increasing energies converging to the ionization potentials (IPs) of the molecule, named Rydberg states, will also appear at low energies in the gas-phase absorption spectrum. To represent such states, it has been employed, as previously done [15, 210], a specific atomic-type one-electron basis set of diffuse character placed on the molecular centroid. The lowest Rydberg series will be represented by excitations (basically single excitations) from the HOMO-like orbital to each of the orbitals of the $n = 3$ series, $3s3p3d$, where n has a value one unit more than the valence main quantum number. As the required valence $\pi\pi^*$ active space is small, previous studies at the CASPT2 level [15, 210] employed an active space of two electrons in 11 orbitals, including the two valence $\pi\pi^*$ plus the nine $3s3p3d$ Rydberg orbitals. As it was soon detected in polyenes [210], the CASSCF procedure is unable to deal properly with the simultaneous calculation of valence and Rydberg states. The lack of correlation leads to wave functions in which the MOs are strongly mixed - the so-called valence-Rydberg mixing -, yielding too diffuse valence states and too compact Rydberg orbitals that only the multi-state CASPT2 is able to correct. Compared with these previous calculations, the RASPT2 results on Table 6.3 can answer the following questions. First, is there any simple partition of the active space that avoids the costly inclusion of the Rydberg orbitals within the CAS space? Second, what is the origin of the valence-Rydberg mixing, and how does RASPT2 handle this problem? Finally, is the multi-state treatment still needed and is there any affordable additional solution?

Table 6.3: Excitation energies (eV) of selected states of Ethene (D_{2h})[a].

State	CASPT2 (2,11)[b]	MS-CASPT2 (2,11)[b]	RASPT2[c] (2,0,1;0,2,9) (S)	MS-RASPT2[c] (2,0,1;0,2,9) (S)	RASPT2[d] (12,3,3;5,2,16) (SDT)	MS-RASPT2[d] (12,3,3;5,2,16) (SDT)	Exp[e]
$1^1B_{1u}(\pi\pi^*)$	8.43	8.04	8.44	8.13	8.06	8.00	8.0[f]
$2^1B_{1u}(3d\pi)$	8.98	9.38	9.03	9.35	9.35	9.41	9.33

[a]Comparison of CASPT2, MS-CASPT2, RASPT2, and MS-RASPT2 results with $\pi\pi^*$ plus Rydberg and $\pi\pi^*$, $\sigma\sigma^*$ plus Rydberg active spaces. [b]CASPT2 and MS-CASPT2, from a State-Average of two $^1B_{1u}$ roots and a 2-electron/11-MO CAS including the two $\pi\pi^*$ MOs and nine ($n = 3$) Rydberg MOs. [c]RASPT2 and MS-RASPT2, from a State-Average of two $^1B_{1u}$ roots and a 2-electron/11-MO RAS including the two $\pi\pi^*$ MOs (in RAS2) and nine ($n = 3$) Rydberg MOs (in RAS3). Only one particle is allowed (S excitations) in RAS3. [d]RASPT2 and MS-RASPT2, from a State-Average of three $^1B_{1u}$ roots and a 12-electron/16-MO RAS including the two $\pi\pi^*$ (in RAS2) MOs, five $\sigma\sigma^*$ MOs (RAS1 and RAS3), nine ($n = 3$) Rydberg MOs (in RAS3), and two additional σ^* MOs (in RAS3). [e]Experimental data. See Ref. [210]. [f]Estimated vertical excitation energy from earlier theoretical work. See therein Ref. [15].

As observed in Table 6.3, and in previous studies [210], for the lowest-energy $^1B_{1u}$ states the perturbative CASPT2(2,11) correction produces values off by almost 0.5 eV compared to experiment. The analysis of the orbital extension $< r^2 >$ (see discussion elsewhere) [15] indicates that even when for the lowest-energy state it should reflect its valence and compact character, yielding a similar value to that for the ground state (1^1A_g), the magnitude for the orbital extension is almost four times larger for both excited states, an illustration of the mixed character of the obtained wave function. It was already proven [15] that the use of the MS-CASPT2 level of calculation is required to get a correct result for the interacting $^1B_{1u}$ states and solve the so-called valence-Rydberg mixing problem. After the orthogonalization produced by the MS treatment the valence and Rydberg states are clearly separated and the corresponding orbital extension - computed by using the Perturbatively-Modified CAS-CI (PMCAS-CI) wave function obtained from the MS method - decreases close to the ground state value for the valence 1^1B_{1u} state, whereas largely increases for the Rydberg 2^1B_{1u} state.

In the RASPT2 calculations, two types of computational strategies have been carried out. First, the nine Rydberg MOs have been placed into the RAS3 active space, leaving the RAS1 space empty and the two $\pi\pi^*$ valence MOs and electrons in RAS2. Only combined singly excited configurations have been allowed toward RAS3, since the Rydberg states are typically well described just by single one-electron promotions, as previously suggested [211]. The active space employed can be labeled as RASPT2(2,0,1;0,2,9)(S), including two active electrons and 11 MOs. Table 6.3 shows that at such level of calculation, that is, by moving the Rydberg MOs to RAS3, there is no loss of accuracy compared to CASPT2(2,11)/MS-CASPT2(2,11). In this small system, the computational effort is only marginally decreased, but the gain will be much more important in larger molecules. Furthermore, two more advantages can be highlighted: additional valence MOs can eventually be added to the RAS2 if required for larger systems, and a single partition of the active space is made available, simplifying considerably the calculations. On the other hand, the behavior of CASPT2(2,11) and RASPT2(2,0,1;0,2,9)(S) with respect to the valence-Rydberg mixing problem is basically the same, as could be expected

by the fact that the same type of correlation effects ($\pi\pi^*$ and Rydberg in both cases) are included in the wave function. Still the MS treatment at the MS-RASPT2 level is required to get correct energies and MOs extensions.

RASPT2 allows enlarging the active space with additional MOs, for instance the σ valence space (5σ, $5\sigma^*$), which were incorporated into RAS1 and RAS3 spaces in the calculations labeled RASPT2(12,3,3;5,2,16)(SDT) in Table 6.3. New states, such as $\sigma\pi^*, \pi\sigma^*, \sigma\sigma^*$ or σRyd* can now be described by this method, but not only that. At the RASPT2(12,3,3;5,2,16)(SDT) level the valence-Rydberg mixing is already solved and the multi-state treatment is not required. As observed, two additional σ^* MOs were finally added to the RAS3 space in order to avoid large intruder state problems. Apart from that, the inclusion of the remaining valence electrons and orbitals in the active space was sufficient to provide an improved wave function and final results within 0.05 eV from experiment. This type of behavior has been observed before when active spaces are enlarged to include correlation effects between MOs of different angular moments [16]. Notice that neither the Rydberg nor the $\sigma\sigma^*$ MOs are included in RAS2. Therefore, in order to incorporate simultaneously the effects of these MOs, a SD excitation level was insufficient (leading to deviations larger than 2 eV) because of the lack of balance between the ground and excited states, as shown in the previous section for FBP. It can be used one strategy which worked for FBP to balance the treatment, namely, the increasement of the level of excitation to SDT. Another option would have been to put all the MOs in RAS2, but this would have been unaffordable in this case. It can be concluded that RASPT2 provides two different solutions to the valence-Rydberg mixing problem, either reaching the MS level of calculation or introducing new MOs into the RAS spaces, if possible.

Table 6.4 presents a comparison between the CASSCF/CASPT2/MS-CASPT2 (2,11) and RASSCF/RASPT2/MS-RASPT2(2,0,1;0,2,9)(S) levels of calculation for the low-lying singlet and triplet valence and Rydberg states in ethene. As in the previous cases, moving the Rydberg orbitals from RAS2 to RAS3, including up to single excitations from RAS1 to RAS3, provides the same type of accuracy than the full inclusion of the Rydberg MOs into RAS2. This recipe

Table 6.4: Excitation energies (eV) of the singlet and triplet valence $\pi\pi^*$ and $n = 3$ Rydberg states of Ethene (D_{2h}).

	CAS(2,11)[a]			RAS(2,0,1;0,2,9)(S)[b]			
State	CASSCF	CASPT2	MS-CASPT2	RASSCF	RASPT2	MS-RASPT2	Exp[c]
1^1A_g							
1^1B_{3u} (3s)	6.57	7.26	7.26	6.45	7.23	7.23	7.11
1^1B_{1g} (3pσ)	7.17	7.91	7.91	7.05	7.88	7.88	7.80
1^1B_{2g} (3pσ)	7.18	7.91	7.91	7.06	7.89	7.89	7.90
1^1B_{1u} (V)	7.93	8.43	8.05	7.83	8.44	8.13	8.0[d]
2^1A_g (3pπ)	7.83	8.31	8.31	7.72	8.26	8.27	8.28
2^1B_{3u} (3dσ)	8.01	8.81	8.81	7.88	8.78	8.78	8.62
3^1B_{3u} (3dδ)	8.11	8.93	8.93	7.98	8.90	8.90	8.90
1^1B_{2u} (3dδ)	8.11	8.96	8.96	7.98	8.94	8.94	9.05
1^1A_u (3dπ)	8.10	8.93	8.93	7.97	8.91	8.91	
2^1B_{1u} (3dπ)	9.38	8.98	9.37	9.37	9.03	9.35	9.33
1^3B_{1u} (V)	4.30	4.44	4.44	4.18	4.41	4.42	4.36
1^3B_{3u} (3s)	6.49	7.17	7.17	6.36	7.15	7.15	6.98
1^3B_{1g} (3pσ)	7.14	7.87	7.87	7.02	7.86	7.86	7.79
1^3B_{2g} (3pσ)	7.15	7.88	7.88	7.02	7.85	7.85	
2^3A_g (3pπ)	7.31	8.17	8.17	7.19	8.11	8.11	8.15
2^3B_{3u} (3dσ)	7.99	8.79	8.80	7.86	8.77	8.78	8.57
3^3B_{3u} (3dδ)	8.05	8.88	8.89	7.92	8.86	8.86	
1^3B_{2u} (3dδ)	8.08	8.94	8.94	7.95	8.92	8.92	
1^3A_u (3dπ)	8.10	8.94	8.94	7.97	8.92	8.92	
2^3B_{1u} (3dπ)	8.41	9.09	9.10	8.28	9.04	9.04	

[a]CASSCF, CASPT2, and MS-CASPT2 results, two electrons and 11 MOs including the two $\pi\pi^*$ MOs and nine ($n = 3$) Rydberg MOs. [b]RASSCF, RASPT2, and MS-RASPT2 results, two electrons - 11 orbitals including the two $\pi\pi*$ MOs (in RAS2) and nine ($n = 3$) Rydberg MOs (in RAS3). Only one particle is allowed (S excitations) in RAS3. [c]See Ref. [210]. [d]Estimated vertical excitation energy from earlier theoretical work. See therein, Ref. [210].

is reliable and a much less costly alternative for the simultaneous calculation of valence and Rydberg states, especially useful for larger systems.

In Table 6.5, excitation energies for benzene are reported at different levels of theory. Valence and Rydberg ($n = 3$ series) singlet excited states of $\pi\pi^*$, $\pi\sigma^*$ and $\sigma\sigma^*$ character calculated with the CASSCF/CASPT2, RASSCF/RASPT2, MS-CASPT2 , MS-RASPT2 and CCSD methods are compiled. Two RASPT2 strategies have been followed. In the first set of calculations, the six $\pi\pi^*$ valence MOs were left in RAS2, and the nine Rydberg orbitals were placed in RAS3 allowing up to single excitations. As in the case of ethene no loss of accuracy is observed with respect to CASPT2 when using this procedure, which largely reduces the computational effort. For instance, the active spaces CAS(6,15) and RAS(6,0,1;0,6,9) generate 2345 and 211 CSFs of ^1A_g symmetry, respectively. In benzene, because of its high symmetry, the spurious mixing of valence and Rydberg wave functions is not such a problem as in ethene, and therefore there is no significant difference when introducing the MS correction, except for symmetries with close-lying valence and Rydberg states like $^1E_{1u}$, where the changes in energies reach up to 0.18 eV.

In the second set of calculations, also reported in Table 6.5, six additional $\sigma\sigma^*$ electrons and MOs have been included, three in RAS1 and three in RAS3, and up to triple excitations have been allowed. These calculations, RASPT2/MS-RASPT2(12,3,3;3,6,12)(SDT), are much more expensive than the previous ones, RASPT2/MS-RASPT2(6,0,1;0,6,9)(S), (e.g., 927588 CFSs for ^1A_g states) and a similar accuracy is obtained. They slightly improve the results in conflictive states like the 2^1E_{2g} valence state, predicted at 7.94 eV at this level, for which the experimental value [212] is 7.8 eV. This highly multiconfigurational state is poorly described by CCSD, which yields 9.18 eV, a value 1.4 eV off from respect to experiment [212]. The inclusion of the $\sigma\sigma^*$ MOs and electrons allows also the computation of new states. As an illustration, the 3^1E_{2g} ($\sigma 3s$) Rydberg state has been computed, a single-reference state, for which RASPT2 and CCSD predict a similar excitation energy [212]. Triple excitations have been included in selected cases in order to compute additional $\sigma\sigma^*$ states. As in the ethene and FBP cases, this is one possible

Table 6.5: Excitation energies (eV) for the low-lying valence and Rydberg singlet states of Benzene (D_{6h})[a].

State	CASPT2[b]	MS-CASPT2[b]	RASPT2[c] (6,0,1;0,6,9) (S)	MS-RASPT2[c,d] (6,0,1;0,6,9) (S)	RASPT2[e] (12,3,3;3,6,12) (SDT)	MS-RASPT2[e,d] (12,3,3;3,6,12) (SDT)	CCSD[f]	Exp[g]
V-ππ*								
1^1B_{2u}	4.94	4.94	4.98	4.98	4.72	4.93	5.19	4.90
1^1B_{1u}	6.22	6.21	6.20	6.20	5.83	6.44	6.59	6.20
1^1E_{1u}	7.12	6.92	7.00	6.82	6.72	6.93	7.17	6.94
2^1E_{2g}	8.05	8.05	8.09	8.10	7.93	7.94	9.18	7.8
R-ππ*								
2^1E_{1u} (3pπ)	7.22	7.29	7.28	7.30	7.16	7.39	7.58	7.41
2^1A_{1g} (3dπ)	7.88	7.87	7.94	7.93	7.85	7.87	7.86	7.81
1^1E_{2g} (3dπ)	7.91	7.88	7.95	7.96	7.88	7.85	7.85	7.81
1^1A_{2g} (3pπ)	7.89	7.91	7.93	7.92	7.82	7.84	7.88	
R-πσ*								
1^1E_{1g} (3s)	6.54	6.54	6.54	6.54	6.50	6.50	6.55	6.33
1^1A_{2u} (3pσ)	7.12	7.12	7.08	7.08	7.10	7.06	6.99	6.93
1^1E_{2u} (3pσ)	7.22	7.22	7.24	7.24	7.18	7.11	7.06	6.95
1^1A_{1u} (3pσ)	7.15	7.15	7.16	7.16	7.11	7.18	7.14	
1^1B_{2g} (3dσ)	7.75	7.79	7.75	7.75	7.67	7.69	7.66	
1^1B_{1g} (3dσ)	7.76	7.79	7.75	7.75	7.68	7.66	7.66	
2^1E_{1g} (3dδ)	7.73	7.72	7.73	7.71	7.69	7.69	7.64	7.54
3^1E_{1g} (3dδ)	7.77	7.76	7.77	7.79	7.71	7.74	7.70	
R-σσ*								
3^1E_{2g} (σ3s)					9.08	9.38	9.39	

[a]For degenerated D_{6h} states two similar values are obtained in D_{2h} in the CASPT2 and RASPT2 steps, unlike in CASSCF or RASSCF where external constraints avoid the orbital mixing and symmetry breaking. In all cases the highest-energy solution has been selected. [b]The CAS space differs for each symmetry. It includes the six valence $\pi\pi^*$ orbitals and those Rydberg orbitals required to obtain the Rydberg states. All energies referred to ground states with the equivalent CAS. [c]RAS1 empty, RAS2 $\pi\pi^*$ valence MOs, and RAS3 including nine ($n = 3$) Rydberg MOs. A single particle (S) allowed in RAS3. [d]A single-root calculation for the 1^1A_g ground state was used in the MS results. [e]Six additional $\sigma\sigma^*$ electrons and MOs added to RAS1 and RAS3, up to three holes/particles allowed in RAS1/RAS3. [f]Linear Response-CCSD calculations. See Ref. [212].[g]See revision of data in Ref. [212].

strategy to compensate for the loss of balance caused by not including in RAS2 the $\sigma\sigma^*$ MOs relevant for the simultaneous description of the ground and excited states. Alternatively, those MOs could be added to RAS2 for both states and then just up to double excitations would be required.

6.1.3. Transition metal compounds and the double d-shell effect: the nickel atom and the copper tetrachloride dianion

Because of the strong correlation effects associated with the $3d$ shell in first-row transition metals (TM), the inclusion of a second correlating d shell ($4d$) in the active space was shown to have crucial effects on the relative state energies obtained from CASPT2 for molecules containing first-row transition metal atoms with a more than half-filled $3d$ shell [6, 11, 213, 214, 215]. This effect, referred to as the double d-shell effect, is manifested in particular when dealing with transitions between states with a different $3d$ occupation number, e.g. $3d \rightarrow 4s$ transitions or charge-transfer (CT) transitions. The double d-shell effect was first reported in a CASPT2 study of the low-lying states of the nickel atom [213]. Here, the results of a comparative CASPT2/RASPT2 study of the lowest states in the electronic spectra of the nickel atom and of the copper tetrachloride dianion ($CuCl_4^{2-}$) are reported. The underlying motivation of this study is to check whether it might be possible to treat the double d-shell effect by means of the much cheaper RASPT2 strategy by, for instance, moving the $4d$ shell into RAS3. This would allow the applicability of the present multiconfigurational approach to more extended and complex TM systems that have so far been inaccessible, or could only be treated qualitatively, because of size limitations of the CASSCF active space, e.g., systems with multiple TM centers [216].

The calculated results obtained for the spectrum of the Ni atom are presented in Table 6.6 and compared to experiment. The first two columns show the CASPT2 relative energies obtained with an active space containing 10 electrons in either the minimum valence active space ($3d$, $4s$) or extended with

Table 6.6: Comparison between CASPT2 and RASPT2 excitation energies (eV) of the excited states of the Nickel atom.

	CASPT2 3d4s[a]	CASPT2 3d4s4d[a]	RASPT2(10,0,m;0,6,5) (3s3p)3d4s4d[a]			Exp[b]
States			SD	SDT	SDTQ	
^{3}D ($3d^9 4s^1$)	0.00	0.00	0.00	0.00	0.00	0.00
^{3}F ($3d^8 4s^2$)	-0.33	-0.10	0.01	-0.03	0.04	0.03
^{1}D ($3d^9 4s^1$)	0.01	0.28	0.29	0.24	0.24	0.33
^{1}D ($3d^8 4s^2$)	1.16	1.45	1.52	1.52	1.61	1.59
^{1}S ($3d^{10}$)	-0.97	1.79	2.16	2.21	2.01	1.74
^{3}P ($3d^8 4s^2$)	1.48	1.68	1.79	1.76	1.85	1.86
^{1}G ($3d^8 4s^2$)	2.33	2.54	2.63	2.62	2.71	2.65

[a]The RAS partition is $3d4s$ in RAS2 and $4d$ in RAS3, with RAS1 empty, RAS(10,0,m;0,6,5), and different excitation levels in RAS3. Core-valence correlation is computed at the perturbative level ($3s3p$ electrons). CASPT2 and RASPT2(10,0,m;0,6,5) with 18 electrons correlated, basis set ANO-RCC $7s6p4d3f2g$, State-Average of 15 roots for triplet states and State-Average of 19 roots for singlet states except ^{1}S($3d^{10}$), a single root calculation. [b]Experimental data. NIST (National Institute of Standards and Technology) [217].

an extra d shell ($3d$, $4s$, $4d$). The double d-shell effect is clearly illustrated by these results. The CASPT2 excitation energies obtained without a second d shell in the active space strongly deviate from the experimental data 0.3-0.5 eV for all states except ^{1}S($3d^{10}$), for which an exceptionally large error of as much as 2.7 eV is found. After including the second d-shell, the errors are reduced to 0.2 eV for all calculated states. These results might be further improved by including also the ($4p$) shell into the active space, and by further extending the basis set.

The next three columns in Table 6.6 give the results obtained from the performance of RASPT2(10,0,m;0,6,5) calculations, with m being the electrons allowed in RAS3, representing a maximum excitation level from two (SD) up to four (SDTQ). Because of the poor convergence of the RASSCF orbital optimization, the RASSCF(SDT) and RASSCF(SDTQ) energies have been calculated at the CI level without orbital optimization, and using the molecu-

lar orbitals converged at the RASSCF(SD) level. As one can see, even at the SD level the double-shell effect is described reasonably well for most states, the results deviating by at most 0.1 eV with respect to the full CASPT2 results. Minor oscillations are observed when increasing the excitation level to SDT and further to SDTQ, but in general there is no clear sign of a systematic improvement. An exception is again the $^1S(3d^{10})$ state. Here, going from CASPT2 to RASPT2 leads to a significant deterioration of the results, by 0.37 eV at the SD level, and decreasing to 0.22 eV at the SDTQ level. However, it is clear that for this state the RASPT2 description of the double d-shell effect is not converged with respect to the excitation level, and higher levels of excitations are necessary for obtaining quantitative accuracy.

As a final set of calculations, the study at the CASPT2 and RASPT2 levels of the excitation energies of the ligand field (LF) states and a charge transfer (CT) state in the electronic spectrum of the copper tetrachloride dianion ($CuCl_4^{2-}$) has been carried out. In order to compare the results with previous reports [206, 218], the same geometry (planar, D_{4h} , with the Cl ligands on the x, y axes) and basis sets have been used. The ground state (GS), 1^2B_{1g}, has a singly occupied molecular orbital (SOMO), σ-antibonding with predominant Cu $3d_{x^2-y^2}$ character, and the lowest part in the spectrum is built from excitations of an electron out each of the other four $3d$ orbitals, giving rise to three ligand field (LF) states: 1^2B_{2g}, 1^2E_g, and 1^2A_{1g}. An important charge-transfer (CT) state, 2^2B_{1g}, corresponding to an excitation out of the bonding counterpart of the ground state SOMO, is also included in the calculations. This CT state belongs to the same symmetry representation as the ground state, and it was shown previously [1] that the interaction between both states resulting from a MS-CASPT2 treatment gives rise to a strongly enhanced covalent character of the GS Cu-Cl σ bonds, by increasing the chlorine $2p\sigma$ contribution in the GS b_{1g}^* SOMO. The purpose of the present study is therefore not only to investigate whether the electronic spectrum of $CuCl_4^{2-}$ may be satisfactorily reproduced by means of a RASPT2 rather than a CASPT2 treatment, but also to see whether the same covalency enhancing effect for the GS may be obtained from a MS-RASPT2 treatment. The latter may be evaluated by comparing the Mulliken spin populations from the

CASSCF and perturbed modified (PM) CASSCF GS wave functions obtained before and after the multi-state treatment, respectively.

The CASPT2 calculations are based on an active space of 11 orbitals, consisting of the Cu 3*d* and 4*d* shells together with the bonding b_{1g} orbital. In the RASPT2 calculations, the correlating 4*d* shell was transferred into RAS3, leaving RAS1 empty and the other six orbitals in RAS2. This then gives results of the type RASPT2(11,0,*m*;0,6,5), with *m* representing the RAS2 $\rightarrow$ RAS3 excitation level. The calculated excitation energies obtained from either a single-state (SS) or multi-state (MS) treatment are presented in Table 6.7. Looking at the SS results first for the LF states, the results obtained from RASPT2-SDTQ results are virtually indistinguishable from CASPT2. A deterioration of the results is observed when decreasing the RASSCF excitation level to SDT and further to SD, although the accuracy of the results obtained from the latter treatment, within 0.2 eV, is still acceptable. On the other hand, for the CT states the RASPT2 treatment seems to be more problematic, giving an excitation energy that deviates more from the CASPT2 results as the level of excitation is increased. Only two of the states included in the calculations belong to the same B_{1g} representation. A MS treatment will therefore leave the total energy of the other states unaffected, while stabilizing the 1^2B_{1g} ground state and destabilizing the 2^2B_{1g} CT state. This then gives rise to a calculated MS-CASPT2 spectrum in which all three LF states are raised in energy by the same amount, 0.13 eV, as compared to SS-CASPT2, while the 2^2B_{1g} CT state is raised by twice this amount.

Table 6.7: Excitation energies (eV) of $CuCl_4^{2-}$ computed at the CASPT2(11,11) and RASPT2(11,0,m;0,6,5) levels of calculation compared with the available experimental data.

			SS-RASPT2(11,0,m;0,6,5)			MS-RASPT2(11,0,m;0,6,5)			
States	SS-CASPT2 (11,11)	MS-CASPT2 (11,11)	SD	SDT	SDTQ	SD	SDT	SDTQ	Exp[a]
1^2B_{1g} (GS)									
1^2B_{2g} (LF)	1.52	1.65	1.36	1.48	1.52	1.51	1.63	1.67	1.55
1^2E_g (LF)	1.77	1.90	1.60	1.72	1.76	1.75	1.87	1.91	1.76
1^2A_{1g} (LF)	2.00	2.13	1.91	1.92	1.99	2.06	2.07	2.14	
2^2B_{1g} (CT)	4.60	4.86	4.51	4.52	4.42	4.81	4.81	4.73	

[a]See refs. [206, 218]

The results obtained from MS-RASPT2 follow the same trend with respect to SS-RASPT2. As such, the same conclusions concerning the accuracy obtained from RASPT2 for the LF and CT states may be drawn from Table 6.7 as already noted for the SS results. As compared to the experimental excitation energies for the 1^2B_{2g} and 1^2E_g (LF) states, the SS treatment yields better excitation energies than MS-CASPT2. The addition of the MS step does not increase the accuracy of the results at any of the levels, CASPT2 or RASPT2. This is not unexpected because the active space requirements with MS are larger than for the lower-level methods. It has been shown before that the addition of angular correlation, that is, the inclusion of orbitals with different angular momentum quantum numbers in the active space, largely improves the MS results [16].

The most important conclusion to be drawn from the results obtained in this subsection is that, in general, moving the $4d$ shell into the RAS3 space is a good strategy that leads to much less expensive calculations in transition metal systems without considerable loss in accuracy. This then allows for the extension of the methodology to larger systems, both increasing the number of transition metal atoms or including additional ligand molecules. For instance, the number of CSFs in $CuCl_4^{2-}$ computations decreases from near 98000 in a CASSCF(10,11) calculation to 4300, 19000, and 47500 at the RASSCF(SD), SDT, and SDTQ levels, respectively.

Finally, the main remarks of the present RASPT2 study for the organic and inorganic systems of Fig. 6.1 are summarized in chapter 7.

6.2. The Hubbard model for Conjugated and Aromatic Hydrocarbons

The Hubbard model was introduced in the 1960's by Hubbard [20] as the simplest model able to track the transition from independent electron (metallic) to correlated electron (localized atomic) behavior. It continues to play a prominent role in the physics literature as a model for strongly interacting fermions. In particular, it has been the model of choice for understanding high-temperature superconductivity [219] and is also widely employed in studies of conducting polymers [220] among many other applications. More recently, it has been combined with density functional theory (the so-called DFT+U method [221]) to study (anti-)ferromagnetic solids [222] and molecular magnetism [223, 224, 225].

The Hubbard model can serve as an important conceptual tool in chemistry, as perhaps first emphasized by Matsen [226], because it links two classic approaches to chemical bonding. As developed here, it represents an extension of the Hückel model which goes to the Hückel limit as the U^{eff} (See section 1.2) parameters become small. On the other hand, in the large U^{eff} limit the spectrum of the Hubbard Hamiltonian can be mapped onto that of a Heisenberg spin Hamiltonian. At lowest order, the Heisenberg Hamiltonian is formally equivalent to the nearest-neighbor Valence Bond model. Thus, as mentioned in section 1.2, the Hubbard model goes continuously from the molecular orbital to the valence bond limit as the strength of the effective electron-electron interaction is varied, and can be used to link and compare these two fundamental pictures of chemical bonding.

The Hubbard Model, which is widely used in physics but is mostly unfamiliar to chemists, provides an attractive yet simple model for chemistry beyond the self consistent field molecular orbital approximation. The Hubbard Model adds an effective electron-electron repulsion U^{eff} when two electrons occupy the same atomic orbital to the familiar Hückel Hamiltonian. Thus it breaks

the degeneracy between excited singlet and triplet states and allows an explicit treatment of electron correlation as explained in section 1.2. In this work, the parameters of the model, $t, U^{eff}, V,$ andV_c, have been evaluated from high-level *ab initio* calculations on two-site fragments and then to transfer the parameters to large molecules and polymers where accurate *ab initio* calculations are difficult or impossible. The recently developed MS-RASPT2 method has been used to generate accurate potential energy curves for ethene as a function of carbon-carbon bond length, which are used to parameterize the model for conjugated hydrocarbons. Test applications to several conjugated/aromatic molecules show that, even though the model is very simple, it is capable of reasonably accurate predictions for bond lengths and predicts molecular excitation energies in reasonable agreement with those from the MS-RASPT2 method.

In section 1.2, the mathematical form of the model has been introduced to determine the model parameters from accurate potential energy curves for a two-site fragment. The present section is divided into three blocks. First, the potential energy curves for ethene as a function of the carbon-carbon distance have been determined by using *ab initio* methods. The next subsection discusses the mathematical forms used to fit the model parameters to the *ab initio* data, and gives the numerical results for the parameters. And finally, the model has been validated in the prediction of both the ground state geometry and the low-lying excitation spectrum of several other conjugated molecules not used in the parameterization when comparing with the results obtained from more expensive *ab initio* computations.

6.2.1. Computation of potential energy curves for stretched ethene

The electronic spectrum of ethene is dominated by the intense and broad band N-V $\pi \rightarrow \pi^*$ valence electronic transition with the superimposed Rydberg excitation $\pi \rightarrow 3s$ [227]. The maximum of this band occurs at 7.66 eV but

does not correspond to the vertical transition [228]. Strong valence-Rydberg mixing is not common for the lower excited states of medium-sized systems but it frequently occurs for small molecules where the lowest valence electronic states fall in the same region as the onset of the first Rydberg series. Some accurate theoretical studies have led to a final estimate of about 8.0 eV for the vertical transition of the ethene V state [229, 230, 231] . At the ground state geometry, the V state has a mixed valence-Rydberg character which strongly depends on the level of theory used to describe it.

The quantum-chemical multiconfigurational methodology employed to compute the potential energy curves for ethene (D_{2h}) is the RASSCF/RASPT2 method, which has proved its accuracy to estimate excitation energies for organic and inorganic systems [6, 232]. The multi-state (MS) treatment has been used in all calculations, i.e. MS-RASPT2. The large atomic natural orbital ANO-L basis set [203] contracted to $[6s5p4d2f]$ for carbon atoms and $[3s2p1d]$ for hydrogen atoms has been employed. In order to describe the Rydberg states, a set of even-tempered $1s1p1d$ functions has been added on carbon atoms. In the current MOLCAS code, the default IPEA shift [14] of 0.25 au and Cholesky decomposition [184] of the two-electron integrals set to the default threshold of 10^{-4} au have been used. The experimental gas-phase equilibrium geometry of the molecule has been employed [205].

The MS-RASPT2 potential energy curves of the three lowest valence excited states of ethene have been computed as a function of the carbon-carbon distance keeping the carbon-hydrogen bond lengths and bond angles fixed at the experimental data. The potential energy curves for the low-lying excited states are displayed graphically in Fig. 6.2. The active space used is comprised by four σ orbitals placed in RAS1, six orbitals in RAS2 (the π and its correlating π^* orbital plus two more of b_{2g} symmetry, one of them being the $3d_{xz}$ Rydberg orbital, and the two (σ, σ^*) molecular orbitals (MOs) of the carbon-carbon bond), and a total of eight orbitals in RAS3 (corresponding to the antibonding counterparts of the four σ MOs plus extra correlating orbitals in order to minimize the appearance of intruder states resulting from the use of an extended basis set [233, 234]). Twelve electrons are active and

up to two holes/particles are allowed in RAS1/RAS3. Using the notation introduced in section 6.1 the level of calculation of the present work is labelled as MS-RASPT2(12,2,2;4,6,8)(SD), where within parentheses the number of active electrons (12), number of holes (2) and particles (2), as well as the number of active orbitals, 4, 6 and 8 for RAS1, RAS2 and RAS3, respectively, are specified.

The valence-Rydberg mixing problem involving the 1^1B_{1u} (V) and 2^1B_{1u} (3dπ) states was treated by previous studies at the ground-state equilibrium geometry [15, 210]. In the present contribution, this problem is solved with a low cost of computation and without loss of accuracy. Thus, at the gas-phase equilibrium geometry [205], the 1^1B_{1u} (V) and 2^1B_{1u} ($3d\pi$) states are computed vertically at 8.03 and 9.33 eV, respectively, in agreement with the best available results [15, 210, 229, 230, 231, 235]. The chosen strategy is based on the study of free base porphin [232], where the most relevant orbitals are placed in RAS2 and up to doubly excited configurations are allowed in RAS1 and RAS3 subspaces. As the extended RASSCF method includes only a small fraction of the correlation energy, the ionic V state (in the valence bond sense) is placed higher in energy where it may come close to the singlet Rydberg state of the same symmetry, leading to erratic valence-Rydberg mixing as it occurs at the CASSCF level [15]. Normally, as discussed in detail elsewhere [232], when both π and σ correlations are included in the CI reference space at the RASPT2(SD) level, the multi-state procedure is required in order to treat properly the valence-Rydberg mixing.

The potential energies displayed in Fig. 6.2 are not true geometry optimizations, since the C-H bond lengths and bond angles have been held fixed at their experimental values. Both experimental analysis [236], using a semitheoretical method, and theoretical analysis [146] with extrapolation to the basis set and full configuration interaction limits yield $R_e = 1.3305$ Å for the carbon-carbon bond length in ethene, while a conventional MP2/cc-pVTZ calculation [108, 139] gives $R_e = 1.3320$ Å. By contrast, a fit of a quadratic to the ground state minimum calculated here gives a minimum energy of E = -78.42617494 au at 1.3390 Å. Investigation shows that the position of the minimum changes

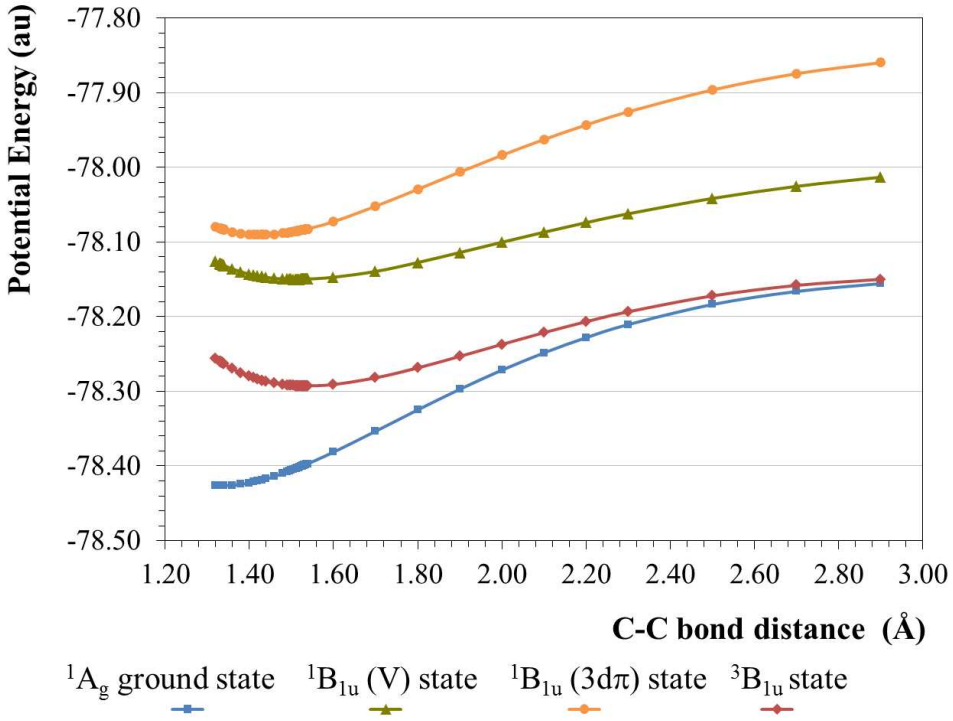


Figure 6.2: Potential energy curves of ethene for the three lowest valence states and the lowest Rydberg state: 1^1A_{1g} (N) (in blue squares), 1^3B_{1u} (T) (in red diamonds), 1^1B_{1u} (V) (in green triangles), and 2^1B_{1u} ($3d\pi$) (in orange circles) performed with RASSCF/MS-RASPT2(12,2,2;4,6,8)(SD) using ANO-L $6s5p4d2f/3s2p1d$ basis set and a set of even-tempered $1s1p1d$ functions on C atoms, and IPEA=0.25 au.

only slightly when the locations of the hydrogen atoms are optimized, so the discrepancy is apparently due to the description of the carbon-carbon bond. The RASPT2 method is expected to give bond lengths which converge (slowly) to the correct value from above.

The calculated vertical excitation energies, at the computed equilibrium carbon-carbon bond length 1.339 Å, of 4.44 eV to the triplet and 8.03 eV to the excited singlet state are in agreement with the available experimental and theoretical data. Experimentally, the triplet is found at 4.36 eV from electron impact [237] and 4.60 eV from optical studies [238], while a variety of accurate calculations fall near 4.50 eV [239]. As mentioned above, the position of the excited valence singlet is not well determined experimentally because of the strong mixing with the nearby Rydberg levels, but the value found here is in agreement with other theoretical calculations [229, 230, 231].

6.2.2. Fitting of the model parameters

The values of the $t, U^{eff}, V,$ andV_c parameters of the Extended Hubbard model, derived from the computed energies displayed in Fig. 6.2, are shown in Table 6.8. The t values found here are close to those found previously [23], but the U^{eff} values are somewhat larger. For transferability to other systems these must be fit to a reasonably simple yet accurate mathematical functional form.

The t parameter

Most semiempirical applications of the Hubbard model take t to be a linear function of bond length. While this is certainly valid over small enough regions, over the full range of bond lengths encountered in conjugated molecules, i.e. $1.32-1.48$ Å, there is noticeable curvature, as shown in Fig. 6.3. Accordingly, t has been initially represented as a power series expansion around the ground state minimum at 1.339 Å. In order to ensure that the model accurately reproduces the ground state minimum, the constant term has been fixed at the *ab intio* t value for 1.339 Å, and the coefficients of higher powers in the expansion have been determined by a least squares fit to the t values at points spaced every 0.02 Å from 1.32 Å to 1.48 Å. The expansion has been then

Table 6.8: *Ab initio* derived Hubbard parameters (in eV)[a].

R (Å)	t	U [eff]	V [b]	V_c [c]
1.320	-3.073557	3.532390	1.467610	1.009559
1.330	-3.027331	3.560306	1.439694	0.902497
1.335	-3.004387	3.574258	1.425742	0.851548
1.339	-2.986420	3.584430	1.415570	0.812838
1.340	-2.981622	3.586978	1.413022	0.802914
1.360	-2.891871	3.636231	1.363769	0.626118
1.380	-2.804717	3.683133	1.316867	0.476268
1.400	-2.719363	3.724482	1.275518	0.350708
1.420	-2.636059	3.761006	1.238994	0.247524
1.440	-2.554715	3.793465	1.206535	0.164409
1.460	-2.475355	3.822664	1.177336	0.099393
1.480	-2.398358	3.847285	1.152715	0.052018
1.500	-2.323576	3.867911	1.132089	0.020384
1.520	-2.251352	3.883726	1.116274	0.004011
1.536	-2.195000	3.894015	1.105985	0.000000

[a]27.211338386 eV = 1 au [b]Assuming $U = 5.0$ eV [c]$V_c(R) = E_T(R) - E_T(R_{eq})$

reexpressed with respect to a more convenient reference point, $R_0 = 1.40$ Å, which is the approximate length of aromatic carbon-carbon bonds as well as the approximate average bond length in linear conjugated polyenes. The final formula for t, in eV, as a function of the bond length R, in Å, is expressed as:

$$t(R) = -2.719288 + 4.22019(R - 1.4) - 2.6072(R - 1.4)^2 \qquad (6.1)$$

Eq. 6.1 gives t values which differ from the *ab initio* values by no more than 0.0004 eV over the range of the fit.

In Table 6.9, it is compared the t values derived from the *ab initio* data and those calculated by using Eq. 6.1.

The V parameter

According to our interpretation of the Extended Hubbard model, the distance dependence of V, shown graphically in Fig. 6.3, should be related to that of U^{eff} by the simple expression:

$$V(R) = U - U^{eff}(R) \tag{6.2}$$

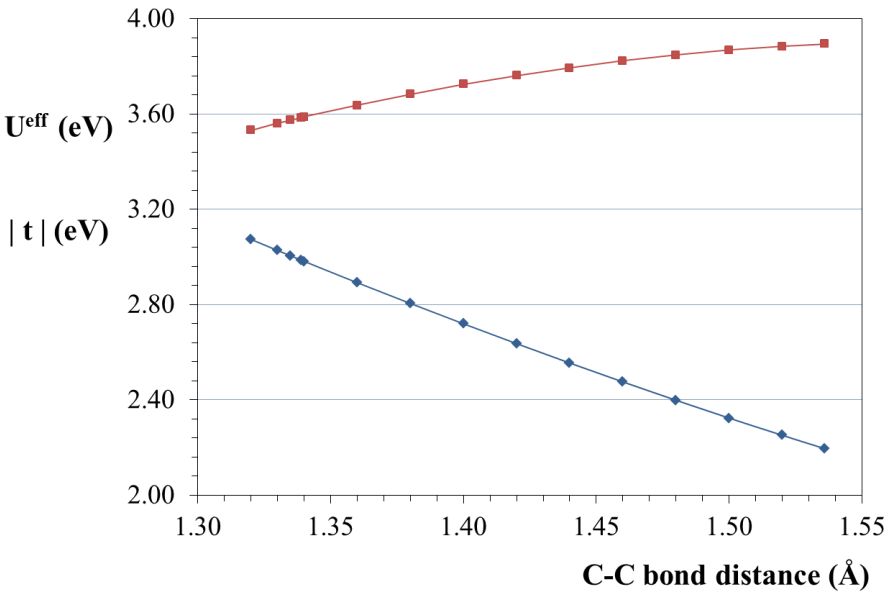


Figure 6.3: The *ab initio* derived values of the Hubbard t parameter (in blue diamonds) and U^{eff} parameter (in red squares) as a function of carbon-carbon bond length.

Table 6.9: Comparison of fits of the t parameter. t values in eV and errors in meV.

R (Å)	t ^{exact}	t (Eq. 6.1)	t error (Eq. 6.1)
1.32	-3.073557	-3.073589	-0.032
1.34	-2.981622	-2.981885	-0.263
1.36	-2.891871	-2.892267	-0.396
1.38	-2.804717	-2.804735	-0.018
1.40	-2.719363	-2.719288	+0.075
1.42	-2.636059	-2.635927	+0.132
1.44	-2.554715	-2.554652	+0.063
1.46	-2.475355	-2.475463	-0.108
1.48	-2.398358	-2.398359	-0.001

The *ab initio* values of U^{eff} have been fit on the same grid of points used for t to a power series expansion around $R_0 = 1.40$ Å. Addition of a quartic term was found to give very little improvement, so the fitting function for U^{eff} is estimated as:

$$U^{eff}(R) = 3.724017 + 1.95894(R - 1.4) - 5.3549(R - 1.4)^2 + 1.457(R - 1.4)^3 \quad (6.3)$$

which, assuming U is fixed at 5.0 eV, gives the following expression for V:

$$V(R) = 1.275983 - 1.95894(R - 1.4) + 5.3549(R - 1.4)^2 - 1.457(R - 1.4)^3 \quad (6.4)$$

being energy given in eV and distance in Å. This expansion is not as accurate as that for t, but V has much less effect on the predictions of the model than t, so it appears to be sufficient.

The V_c parameter

According to the Hubbard model interpretation, the π electrons make no contribution to the energy of the triplet state. Its energy is then just the energy of the underlying σ-bonded core. Thus $E_T(R)$ should represent the energy of the σ bond between the carbon atoms. In fact, it shows a minimum between 1.52 and 1.53 $\mathring{A}$, which agrees well with the expected equilibrium bond length for a σ bond between two sp^2 carbon atoms [240]. In comparison, the corresponding triplet state in the two-electron system of H_2 is purely repulsive since the underlying potential is just the repulsion between the bare nuclei. $V_c(R)$ is modeled as an anharmonic expansion about the minimum of the triplet state, including both a cubic and a quartic anharmonicity.

A fit of a quadratic to the three *ab initio* points nearest the minimum gives a minimum energy of $E_{eq} = -78.29274234$ au at $R_{eq} = 1.5361$ Å. The $V_c(R)$ parameter is then represented as:

$$V_c(R) = E_{eq} + \frac{1}{2}K(R - R_{eq})^2 + \frac{1}{6}K'(R - R_{eq})^3 + \frac{1}{24}K''(R - R_{eq})^4 \quad (6.5)$$

The three force constants K, K', and K'' have been determined by a least squares fit to *ab initio* points spaced at 0.02 Å intervals between 1.32 Å and 1.52 Å. The value of E_{eq} just shifts the energy of all states by a constant and has no influence on any of the predictions of the model. The final expression for $V_c(R)$ is obtained as:

$$V_c(R) = E_{eq} + \frac{1}{2}(31.0140)(R - 1.5361)^2 + \frac{1}{6}(-117.885)(R - 1.5361)^3 + \frac{1}{24}(955.08)(R - 1.5361)^4 \quad (6.6)$$

being energies given in eV and distances in Å. The harmonic force constant of about 31 eV/(Å)^2 is of just the value expected [240, 241] for a carbon-carbon single bond force constant, confirming the interpretation of E_T as essentially the energy of the σ bond in ethene.

6.2.3. Testing and validation of the model

Eq. 6.1, 6.4, and 6.6, along with the fixed value $U = 5.0$ eV, define the parameterization of the Extended Hubbard model. By construction, they accurately model ethene, but a critical test of the soundness of the model is whether the parameters are transferable to other hydrocarbons. Accordingly, the model has been applied to three typical molecules, *trans*-1,3-butadiene which introduces conjugation, *trans*-1,3,5-hexatriene which introduces conjugation extending beyond nearest neighbors, and benzene which shows the much more dramatic effect of conjugation around a cycle of $4n$+2 atoms. For each molecule, the equilibrium carbon-carbon bond lengths have been determined and low-lying electronic excitation energies have been computed.

Bond lengths

Care must be taken in comparing theoretically calculated equilibrium bond lengths R_e (the lowest point on the potential energy curve) with experimentally measured bond lengths. Experiments never measure R_e directly [242],

Table 6.10: Carbon-carbon bond lengths (in Ångström) of ethene, *trans*-1,3-butadiene, *trans*-1,3,5-hexatriene, and benzene at the ground-state equilibrium geometry.

Bond	MP2/cc-pVTZ[a]	MP2/6-31G(d,p)[a]	Extended Hubbard Model	Exp[b]
C_2H_4				
double	1.332	1.335	1.339	1.339
C_4H_6				
double	1.340	1.343	1.350	1.349
single	1.453	1.459	1.461	1.467
C_6H_8				
end	1.342	1.345	1.352	1.337
single	1.446	1.450	1.454	1.458
middle	1.350	1.353	1.362	1.368
C_6H_6				
aromatic	1.394	1.396	1.397	1.399

[a]Calculated with the Gaussian 03 program package [191]. [b]Reference [243].

and experimental bond lengths are often longer than R_e, sometimes by as much as 0.01 Å. For example, the conventionally quoted [243] carbon-carbon bond length in ethene of 1.339 Å, measured by microwave spectroscopy, is substantially longer than the accurately determined [146, 236] value of $R_e = 1.3305$ Å. For diatomic molecules, it is usually possible to determine R_e using only experimentally measured quantities, but for polyatomic molecules only a few accurate experimentally based R_e values are known [236]. Accordingly, it is established a comparison of Hubbard model bond length predictions either with those computed by conventional quantum chemical methods or with experimental measurements for the ground state.

In theoretical calculations it is generally acknowledged that electron correlation must be included to obtain accurate bond lengths, but correlation at the MP2 level is usually regarded as sufficient. Geometry optimizations are often

carried out at the MP2 level using a Pople-style 6-31G(d,p) basis [99], but for carbon-carbon bond lengths it appears that the cc-pVTZ basis may be more accurate [244]. On the other hand, since the MS-RASPT2 bond length for ethene agrees with the experimental ground state value, it may be more relevant to compare the Hubbard model bond lengths to experiment for other small molecules as well. In Table 6.10 bond lengths found by minimizing the energy of the Hubbard model as a function of bond length are compared to those obtained from MP2 calculations with both basis sets, and to experiment.

Inspection of Table 6.10 shows a systematic difference between the two *ab initio* basis sets. The 6-31G(d,p) basis gives longer lengths for all bonds, especially single bonds, and this difference persists in longer polyenes [244]. The Hubbard model bond lengths are generally between the *ab initio* and the experimental values, with the exception of the end bond in hexatriene. With that exception, the agreement between the Hubbard model predictions and experimental bond lengths is quite good, with the largest remaining error just 0.006 Å. Agreement with *ab initio* bond lengths, particularly those calculated with the 6-31G(d,p) basis at the MP2 level, is also quite reasonable, indicating that the model has indeed captured the most important factors determining molecular structure.

Excitation energies

Care must also be taken when comparing theoretically calculated excitation energies to experiment since experiments measure transitions between specific vibrational states. Nevertheless, experimental excitation energies measured at the band maximum often agree well with theoretical vertical transition energies. In Table 6.11, electronic transition energies to low-lying excited states computed from the Hubbard model are compared both to *ab initio* calculations using the extended RASPT2 method [232] and to experiment [239]. Two different results from the model are shown, one at the cc-pVTZ *ab initio* geometry and one at the model equilibrium geometry.

Table 6.11: Excitation energies (in eV) of the lowest valence singlet and triplet states of ethene, *trans*-1,3-butadiene, *trans*-1,3,5-hexatriene, and benzene at the ground-state equilibrium geometry.

Molecule	Exp[a]	MS-RASPT2[b,c]	Ext. Hubbard Model[c]	Ext. Hubbard Model[d]
C_2H_4				
1^1B_{1u}	8.00[e]	8.07	8.08	8.03
1^3B_{1u}	4.36, 4.60	4.51	4.51	4.44
C_4H_6				
1^1B_u	5.92	5.97	5.71[f]	5.65[f]
2^1A_g		6.57	6.37	6.25
1^3B_u	3.20	3.33	3.23	3.15
1^3A_g	5.08	5.12	5.05	4.95
C_6H_8				
1^1B_u	5.13, 4.93	5.01	4.47[f]	4.41[f]
2^1A_g		5.37	5.07	4.95
1^3B_u	2.61	2.68	2.56	2.49
1^3A_g	4.11	4.31	4.12	4.04
C_6H_6				
1^1B_{2u}	4.90	4.91	5.02	4.99
1^1B_{1u}	6.20	6.15	6.34	6.32
1^1E_{1u}	6.94	6.86	7.02	7.00
1^3B_{1u}	3.94	4.09	4.21	4.18
1^3E_{1u}	4.76	4.70	4.98	4.95
1^3B_{2u}	5.60	5.40	6.10	6.08

[a]For original experimental references see Ref. [239]. [b]Present MS-RASPT2(n,0,1;0,n,9)(S) with n the number of π electrons and (π, π^*) MOs of the full valence space. MS-RASPT2(12,2,2;4,6,8)(SD) for ethene. [c]At the MP2/cc-pVTZ geometry. [d] At the Hubbard model geometry. [e]Estimated vertical excitation energy from earlier theoretical work. See Ref. [229, 230, 231] and references therein. [f]Charge transfer states not expected to be well described by the Hubbard model.

The MS-RASPT2 vertical transition energy calculations have been carried out at the *ab initio* equilibrium geometry as described in section 6.1.2, including all $\pi\pi^*$ valence orbitals in RAS2 subspace and the nine $3s3p3d$ Rydberg orbitals in RAS3 subspace, allowing up to singly excited configurations into the latter, RASPT2(n,0,1;0,n,9)(S) with n the number of π electrons, that is equivalent to the number of $\pi\pi^*$ MOs of the full valence space. As in section 6.2.1, the results for ethene have been performed at the RASPT2(12,2,2;4,6,8)(SD) level, employing the $6s5p4d2f/3s2p1d+1s1p1d$ basis set, the default IPEA shift [14], and Cholesky decomposition [184]. An imaginary level shift of 0.1 au has been used throughout. Symmetry restrictions have been imposed for the computed energies in the four molecules, D_{2h} for ethene, C_{2h} for *trans*-1,3-butadiene and *trans*-1,3,5-hexatriene, and D_{6h} (D_{2h} in MOLCAS) for benzene. Ground-state geometry optimizations have been performed at the MP2 level [108] with the polarized cc-pVTZ basis set [139].

As can be seen in Table 6.11, MS-RASPT2 transition energies for ethene are in agreement with accurate multireference CI theoretical studies [229, 230, 231]. The vertical energy for the lowest $^1B_{1u}$ state is computed to be 8.07 eV which is consistent with previous results [229, 230, 231, 232, 235, 239]. For the triplet 1^3B_{1u} state, the MS-RASPT2 vertical transition energy is found at 4.51 eV, falling in the region of the experimental data [238, 245].

As in ethene, valence-Rydberg mixing is a common situation in relatively small organic molecules and it is also present in the next studied polyenes, *trans*-1,3-butadiene and *trans*-1,3,5-hexatriene. The well-known erratic 1^1B_u valence state is well-described by MS-RASPT2 yielding transition energies of 5.97 eV and 5.01 eV for butadiene and hexatriene, respectively. For the 2^1A_g state, the MS-RASPT2 transition energies of both molecules are near to those obtained in previous studies [210, 246]. The two lowest triplet states 1^3B_u and 1^3A_g estimated at the MS-RASPT2 level for both butadiene and hexatriene are close to those obtained by Silva et al. [234] and experimental data [238, 245].

As shown in the past [233], benzene also has valence-Rydberg mixing like the polyenes discussed before. In this aromatic molecule the erratic valence state is 1^1E_{1u}, but MS-RASPT2 solved it properly as shown in section 6.1.2. The three lowest singlet valence states 1^1B_{2u}, 1^1B_{1u}, and 1^1E_{1u} estimated at the MS-RASPT2 level seem to be in agreement with experimental data [239], but the results found from the above-mentioned benchmarks turn out to be somewhat overestimated [234, 239]. For the lowest triplet states 1^3B_{1u}, 1^3E_{1u}, and 1^3B_{2u}, the MS-RASPT2 values are close to the experimental data [239] and results obtained from Ref. [234].

As can be seen from Table 6.11, the MS-RASPT2 results are in agreement with the experimental values. When computed at the same geometry, the Hubbard model results for the lowest excitation of each spin are also in agreement, with two exceptions to be discussed shortly. Errors in higher excited states are somewhat larger, generally in the 0.2-0.3 eV range. The two states for which the Hubbard model gives large errors are the ^1B_u states of butadiene and hexatriene. These are charge transfer states, in which determinants with two π electrons on one end carbon and none on the other end play a significant role. They are not expected to be well described by any nearest-neighbor model which neglects long range electron-electron repulsion, including the Hubbard model. The Hubbard model does agree with CASPT2 and most [239, 247] (but by no means all [248, 249]) *ab initio* calculations in placing the ^1B_u charge transfer states below the ^1A_g homopolar states in these molecules. The homopolar state is spectroscopically dark, but there is no doubt that it becomes the lowest-lying singlet excited state in longer polyenes [250], with the cross-over probably coming at octatetraene in which the two states are nearly degenerate.

Calculations using the Hubbard model at its own equilibrium geometry rather than the *ab initio* geometry show only small shifts of less than 0.1 eV. These shifts generally act to worsen agreement with experiment for butadiene and hexatriene, but to improve the agreement for benzene. Nevertheless, the predictions of the model are quite satisfactory, with all states in the correct order and with the differences with respect to the MS-RASPT2 results no larger

than those produced by other *ab initio* methods [239]. Thus, it seems again that the Hubbard model has identified the important interactions needed to understand electronic spectra.

6.3. Photophysics of the $B_{18}H_{22}$ boron hydrides

The *anti*- and *syn*-isomers of octadecaborane(22) (Fig. 6.4), valued for their high molecular boron content, ready synthesis, and stability in air, are two of the more important binary boron hydrides available for practical applications. In recent years they have received particular attention from the semiconductor industry as new-generation dopant materials used in the manufacture of p-type silicon. As a dopant, octadecaborane(22) makes shallow boron implants into silicon wafers at high throughputs and low energies, which are significant industrial advantages [45]. Running parallel to the development of this technology is the requirement to better understand the chemistry and physics of octadecaborane(22) at fundamental levels.

In this context, there is an interest in the physico-chemical differences between the two isomers of $B_{18}H_{22}$. Both the *syn*- and *anti*-$B_{18}H_{22}$ isomers are structurally related to commercially available decaborane(14). The structure of *syn*-$B_{18}H_{22}$ can be formally derived from the fusion of two B_{10} units which share a common B-B edge so that the molecule has a two-fold axis of symmetry (Fig. 6.4a). In the case of the *anti*-$B_{18}H_{22}$ isomer, the two halves of the molecule are related by a centre of symmetry (Fig. 6.4b).

Hawthorne, to whom much of the early work on the chemistry of octadecaborane(22) can be ascribed, recorded the absorption spectra of *syn*- and *anti*-$B_{18}H_{22}$ under UV-vis irradiation [39]. However, besides the simple documentation of the absorption bands, useful for compound identification, and the note that *solutions of the hydride in hydrocarbon solvents exhibit a purple fluorescence* (no emission spectra were reported), there has been no further consideration of this interesting property. This is somewhat surprising as visible fluorescence is a unique property in the neutral binary boron hydrides [251]. Furthermore, the fluorescence is observed only in *anti*-$B_{18}H_{22}$ and is absent in similar solutions of *syn*-$B_{18}H_{22}$. This interesting difference between the two isomers is deserving of closer investigation, especially due to the in-

terest in the physico-chemical differences between the two isomers of $B_{18}H_{22}$. Thus, in order to gain a deeper insight into the electronic structure of the two isomers and to find an explanation for the distinctive fluorescence behavior of the isomers, a UV-vis spectroscopic study of both species has been carried out using contemporary time-resolved techniques, which are supported and supplemented by the employment of sophisticated computational methods.

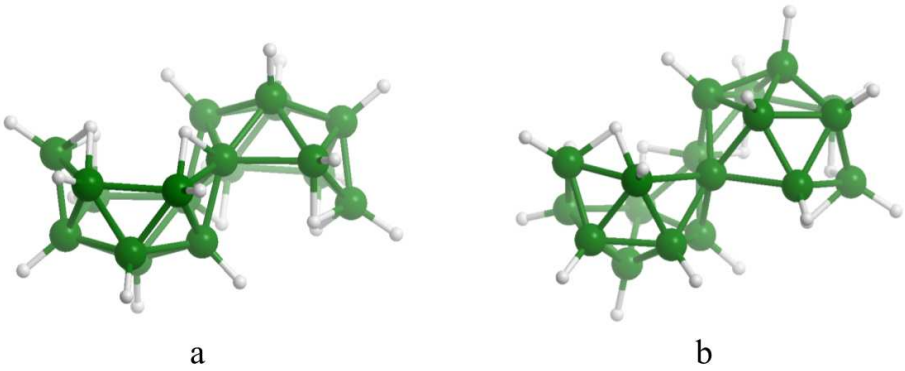


Figure 6.4: Molecular structures of *anti*-$B_{18}H_{22}$ (a) and *syn*-$B_{18}H_{22}$ (b) isomers.

The present section is divided into three subsections. First, the experimental and theoretical details are introduced. Next, the experimental results are discussed and, finally, the theoretical calculations are compared with experimental data.

6.3.1. Experimental procedure and theoretical methodology

The two isomers of $B_{18}H_{22}$ have been synthesized from the hydrolysis of the hydronium ion salt of $[B_{20}H_{18}]^{2-}$ according to the method described in the literature [39]. Pure samples of each isomer, obtained by repeated crystalliza-

tions from hot saturated *n*-hexane solutions, had ^1H and ^{11}B NMR and mass spectroscopic characteristics in agreement with the published data [39, 252].

Details of UV-vis spectroscopic and luminescence measurements are described elsewhere [253] together with the setup used to obtain femtosecond transient absorption spectra and time-resolved near-infrared phosphorescence of $O_2(^1\Delta_g)$.

The fluorescence quantum yield of *anti*-$B_{18}H_{22}$, Φ_F, in hexane has been measured by the comparative method with quinine sulfate as a reference (0.1 M H_2SO_4, oxygen-free, $\lambda_{exc} = 350$ nm, $\Phi_F^R = 0.58$) [254]. The quantum yield has been calculated using the following expression:

$$\Phi_F = \Phi_F^R \left(\frac{F}{F^R}\right) \left(\frac{A^R}{A}\right) \left(\frac{n}{n^R}\right)^2 \tag{6.7}$$

where F is the integrated area under the emission spectrum, A is absorbance at the excitation wavelength, and n is the respective refractive index of the solvent. Superscript R corresponds to the reference.

The averaged kinetic traces of the phosphorescence signal of $O_2(^1\Delta_g)$ have been fitted by the following fitting function with τ_T (triplet states lifetime), τ_Δ (lifetime of $O_2(^1\Delta_g)$), and $\mathcal{A}$ (amplitude) as free parameters:

$$S(t) = \mathcal{A}\frac{\tau_\Delta}{\tau_T - \tau_\Delta} \left(exp\left(-t/\tau_T\right) - exp\left(-t/\tau_\Delta\right)\right) \tag{6.8}$$

The quantum yield of singlet oxygen formation, Φ_Δ, in hexane has been estimated by the comparative method using 5,10,15,20-tetraphenylporphyrin (TPP, $\Phi_\Delta = 0.6$ in benzene) [255] as a standard.

In the theoretical calculations, both Density Functional Theory (DFT) and *ab initio* multiconfigurational methods have been used. The complete active space self-consistent field (CASSCF) approach followed by multiconfigurational second-order perturbation theory (CASPT2), have been used to compute potential energy hypersurfaces (PEH) minima and crossings, together with state and transition properties for the lowest-lying singlet and triplet states of both $B_{18}H_{22}$ isomers. The ground (S_0) and low-lying triplet (T_1) states minima have been optimized and characterized at the DFT/B3LYP/6-31+G(d) level of theory. All other calculations, including determinations of PEH minima and crossings, have been performed with the CASSCF method supplemented, at the obtained points, with CASPT2 calculations in order to include the necessary dynamical correlation effects. The CASPT2//CASSCF protocol [9, 10, 109] has proven its accuracy repeatedly [16, 210]. The conical intersection structures have been estimated as minimum energy crossing points (MECP) between the two states involved. The one-electron atomic basis set 6-31+G(d) has been used throughout. The final results use an active space of 12 molecular orbitals and 12 electrons. No symmetry restrictions have been imposed during the MECP calculations. At the ground state, the *anti*-$B_{18}H_{22}$ and *syn*-$B_{18}H_{22}$ isomers have been determined to have C_i and C_2 geometries, respectively, and therefore the corresponding state labeling is used when required. Oscillator strengths and radiative lifetimes have been computed from the CASSCF TDMs and the CASPT2 excitation energies as explained in chapter 5. Spin-orbit coupling (SOC) terms between singlet and triplet states have been estimated within the AMFI and CASSI frameworks, included in the MOLCAS quantum-chemistry package [16, 256].

6.3.2. Experimental Results

The absorption spectrum of *anti*-$B_{18}H_{22}$ in hexane is characterized by broad bands at 272 and 329 nm that are bathochromically shifted to 229 nm and 308 nm for the *syn*-$B_{18}H_{22}$ isomer (Table 6.12, Fig. 6.5). Both isomers differ considerably in their fluorescence properties (Table 6.12). *Anti*-$B_{18}H_{22}$ shows

very strong fluorescence peaking at 407 nm in hexane (Fig. 6.5) with a fluorescence quantum yield, Φ_F, of 0.97 ± 0.02 and a relatively long fluorescence lifetime of (11.2 ± 0.1) ns. The superimposition of the excitation spectra and the absorption spectra serves as an experimental confirmation of the identity of the fluorescence. The energy level of the lowest-lying singlet excited state has been estimated from the intersection of the normalized absorption and emission bands, giving a value of 3.41 eV. In contrast, *syn*-$B_{18}H_{22}$ has no detectable fluorescence emission.

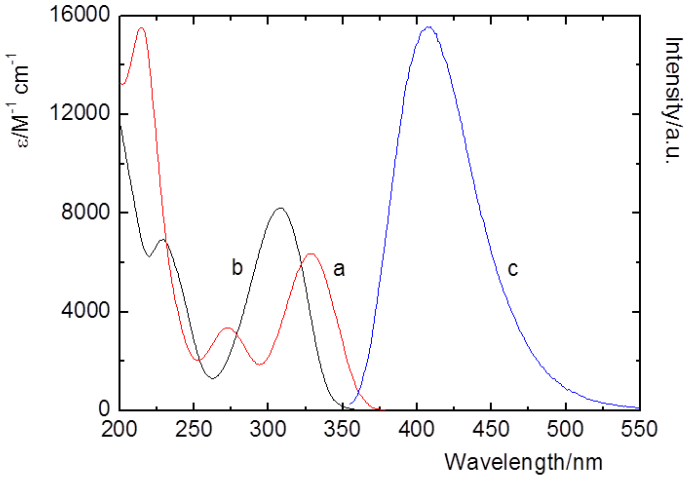


Figure 6.5: UV-vis absorption spectra of *anti*-$B_{18}H_{22}$ (a) and *syn*-$B_{18}H_{22}$ (b) (left axis) and fluorescence emission spectrum (c) (λ_{exc} = 340 nm) of *anti*-$B_{18}H_{22}$ (right axis) in hexane.

Femtosecond transient absorption spectroscopy has been used to detect non-emitting transients produced upon excitation of *syn*-$B_{18}H_{22}$. Femtosecond flash excitation at 330 nm of *syn*-$B_{18}H_{22}$ in hexane affords the absorption profile of the excited singlet state with a maximum at 475 nm (Fig. 6.6b). Fast relaxation to the ground state with a time constant of about 75 ps (Fig. 6.7) is in accordance with the nonfluorescent character of *syn*-$B_{18}H_{22}$. For comparison, the excited singlet states of *anti*-$B_{18}H_{22}$ are characterized by a narrow band at 512 nm and broad absorption above 600 nm (Fig. 6.6a).

Table 6.12: Photophysical properties of *syn*-$B_{18}H_{22}$ and *anti*-$B_{18}H_{22}$ in air-saturated hexane: absorption spectra (molar absorption coefficients ϵ at band maxima λ_{max}), fluorescence band maximum λ_F, quantum yield of fluorescence Φ_F, lifetime of the excited singlet states τ_F, quantum yield the triplet states Φ_T, lifetime of the triplet states τ_T, quantum yield of singlet oxygen formation Φ_Δ.

		Excited singlet states			Triplet states		$O_2(^1\Delta_g)$
$B_{18}H_{22}$ isomer	Absorption spectra λ_{max}/nm (ε/mM^{-1}cm^{-1})	λ_F/nm	Φ_F	τ_F/ns	Φ_T	τ_T/µs	Φ_Δ
syn	229 (7.32), 308 (8.66)		~ 0 [a]	0.075 [b]		< 2[c]	~0.008[c]
anti	215 (15.2), 272 (3.36), 329 (6.29)	407 [a]	0.97±0.02[a]	11.2±0.1[a]	<0.03[a]	6±2[c]	~0.008[c]

[a]Data from fluorescence emission spectroscopy. [b]Data from femtosecond transient absorption spectroscopy. [c]Data from time-resolved near-infrared phosphorescence of $O_2(^1\Delta_g)$.

The kinetics recorded at 512 nm does not decay within the time window of the femtosecond experiment (Fig. 6.7), in agreement with the 11.2 ns lifetime obtained by the fluorescence decay analysis (Table 6.12).

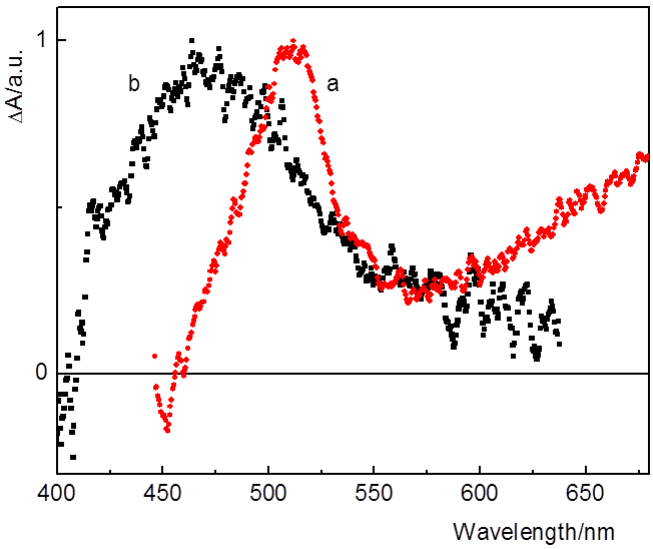


Figure 6.6: Transient absorption spectra of *anti*-$B_{18}H_{22}$ (a) and *syn*-$B_{18}H_{22}$ (b) excited at 330 nm and recorded at 100 ps and 12 ps, respectively.

The observed phosphorescence at 1270 nm, which may be attributed to $O_2(^1\Delta_g)$ (Fig. 6.8, inset), appears after excitation of both isomers, disappears after purging the sample with inert environment, and reappears after the addition of air or oxygen (Fig. 6.9). The steep increase in the $O_2(^1\Delta_g)$ concentration during the first ten μs after excitation of *anti*-$B_{18}H_{22}$ hexane solutions (Fig. 6.8) suggests that $O_2(^1\Delta_g)$ is formed by the diffusion-controlled quenching of the excited states by oxygen, although a contribution from $O_2(^1\Delta_g)$ produced by photodecomposition of a boron cluster-O_2 adduct [257] cannot be excluded. The formation of $O_2(^1\Delta_g)$ from excited singlet states is not expected, since fluorescence of *anti*-$B_{18}H_{22}$ is not quenched by oxygen within experimental error, and *syn*-$B_{18}H_{22}$ has a very short lifetime in its singlet excited states (τ_F

= 75 ps). Therefore, it is suggested that the triplet states are involved in the formation of $O_2(^1\Delta_g)$.

The lifetime of the *anti*-$B_{18}H_{22}$ $O_2(^1\Delta_g)$-producing states, calculated from phosphorescence decay at 1270 nm (Eq. 6.8), is $\tau_T = (6 \pm 2)$ μs in air-saturated hexane (Fig. 6.8). The obtained value of $\tau_\Delta = (32 \pm 4)$ μs is in good agreement with the lifetime of $O_2(^1\Delta_g)$ in pure hexane ($\sim$ 30 μs) [258]. For comparison, the lifetime of these states is longer in chloroform (τ_T = 13 μs), concomitant with the longer lifetime of $O_2(^1\Delta_g)$ (τ_Δ = 170 μs) in chloroform. In the case of *syn*-$B_{18}H_{22}$ the lifetime of the triplet states is found to be shorter than 2 μs. The quantum yield of $O_2(^1\Delta_g)$ formation is estimated to be $\Phi_\Delta \sim 0.008$ for both the *anti*-$B_{18}H_{22}$ and *syn*-$B_{18}H_{22}$ isomers.

Despite several attempts, no transient absorptions have been observed on a microsecond timescale. The fact that the absorption of the $O_2(^1\Delta_g)$-producing

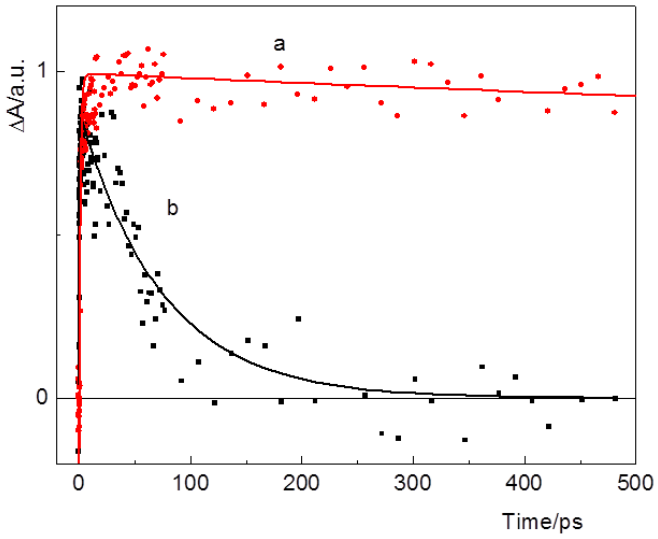


Figure 6.7: Kinetics measured at maxima of transient absorption spectra of *anti*-$B_{18}H_{22}$ (a) and *syn*-$B_{18}H_{22}$ (b) at 512 nm and 475 nm, respectively.

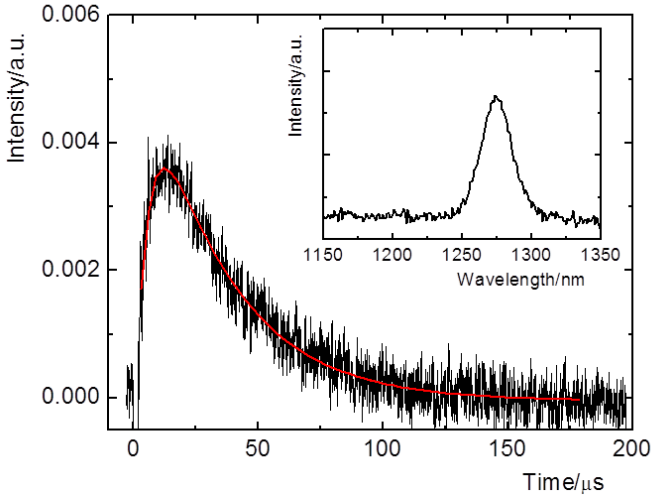


Figure 6.8: Time dependence of the $O_2(^1\Delta_g)$ phosphorescence at 1270 nm produced by *anti*-$B_{18}H_{22}$ ($\lambda_{exc} = 308$ nm) calculated as the difference of the signal in air- and argon-saturated hexane. The smoothed red line is a least squares fit to the pure luminescence signal (Eq. 6.8). Inset: Phosphorescence spectrum of $O_2(^1\Delta_g)$ in an oxygen-saturated chloroform solution of *anti*-$B_{18}H_{22}$ (λ_{exc} = 350 nm).

states and their quenching by molecular oxygen has not been observed can be explained by the very low quantum yield of the triplet state formation (< 0.03) and low triplet(1^3A_u)-triplet oscillator strengths (< 0.1) in the case of *anti*-$B_{18}H_{22}$, and the fast relaxation of the excited states and very low triplet(1^3B)-triplet oscillator strengths (< 0.05) in the case of *syn*-$B_{18}H_{22}$.

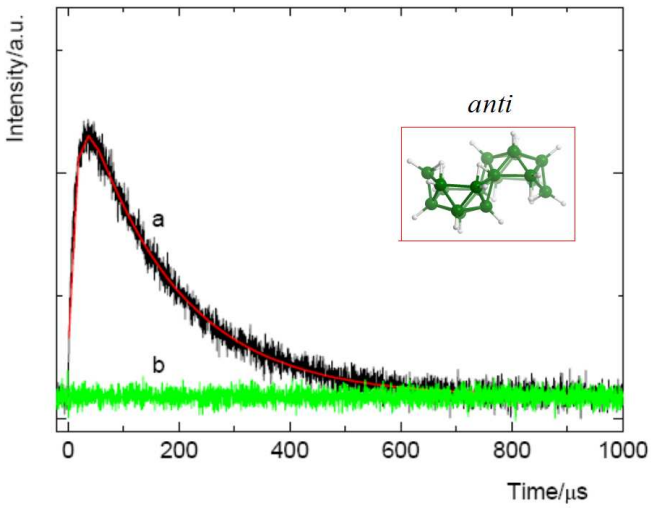


Figure 6.9: Time dependence of the $O_2(^1\Delta_g)$ luminiscence signal at 1275 nm produced by *anti*-$B_{18}H_{22}$ in chloroform (λ_{exc} = 355 nm): air-saturated (a) and N_2-saturated solution (b). The smoothed line (red) is a least squares fit, Eq. 6.8.

6.3.3. Theoretical calculations and comparison with experimental data

Absorption spectra

Selected CASPT2 results and the measured experimental data for the absorption spectrum of *anti*-$B_{18}H_{22}$ are compiled in Table 6.13. The ground state (S_0) minimum geometry, determined under the C_i point group symmetry at the B3LYP level and therefore labeled 1^1A_g, has been employed. Several intense transitions have been computed from the ground state 1^1A_g up to 183 nm (6.76 eV) toward one-photon allowed $1A_u$ states, two of which fall within the boundaries of the recorded experimental data in hexane, up to 200 nm

(6.20 eV). As is typical in many organic systems, the low-lying spectra of the molecule, especially its most intense bands, can be mainly described by the four Platt-type states [259] relating the two highest occupied and two lowest unoccupied molecular orbitals: ^1L_a state, basically described as the HOMO (H) $\rightarrow$ LUMO (L) one-electron promotion; the ^1L_b and ^1B_b states, the respective antisymmetric (low intensity) and symmetric (high intensity) combination of the H-1 $\rightarrow$ L and H $\rightarrow$ L+1 configurations, and the ^1B_a state, described by the H-1 $\rightarrow$ L+1 one-electron promotion (Fig. 6.10). The lowest-energy singlet-singlet transition, $1^1A_g \rightarrow 1^1A_u$, described as the H $\rightarrow$ L one-electron promotion (90%) at the CASSCF level, is computed vertically at 315 nm (3.93 eV) with an oscillator strength of 0.265, and it can be clearly assigned to the band maximum observed for the *anti* isomer at 329 nm (3.77 eV) with an intermediate intensity. Figure 6.10 displays the most relevant CASSCF molecular orbitals for the *anti*-$B_{18}H_{22}$ isomer, which are shown to be distributed along the structure of the borane cage. Next, a transition to the 2^1A_u state is computed at 248 nm (4.99 eV), described essentially as a combination of two configurations, H-1 $\rightarrow$ L (74%) and H $\rightarrow$ L+1 (16%), and with a related oscillator strength of 0.051. As no other intense transition has been computed in the 310-225 nm ($4.0-5.5$ eV) energy interval, this can be unambiguously assigned to the weak and broad band peaking at 272 nm (4.56 eV). It is not surprising that the theoretically computed excitation energies (vertical absorption energies), which are derived from using the ground state optimized geometry, lie higher in energy than the measured band maxima. This is because the equilibrium geometries of the excited states are different from that of the ground state and, with regard to the Franck-Condon principle, the verticality of the transitions is only partially fulfilled.

Apart from the computed weak transition for the 3^1A_u state at 223 nm (5.55 eV), which lies beneath the most intense measured band at 215 nm (5.76 eV), only one other allowed transition has been computed within the range of the experimental spectrum. This is the 4^1A_u state at 212 nm (5.85 eV), which has a large oscillator strength (0.884) and may be assigned to the measured feature at 215 nm (5.76 eV). In turn, this feature can be described as ^1B_b, i.e. the counterpart of the 2^1A_u (^1L_b) transition according to its CASSCF

Table 6.13: Theoretical (CASPT2) absorption and emission spectra of *anti*-$B_{18}H_{22}$.

	CASPT2							
State	E_{VA}[a] /eV	λ_{VA}[a] /nm	f[a]	T_e[a] /eV	λ_e[a] /nm	E_{VE}[a] /eV	λ_{VE}[a] /nm	τ_{rad} / s
1^1A_u	3.93	315	0.265	3.63[c]	341[c]	2.91[c]	426[c]	5.6×10^{-9}
2^1A_u	4.99	248	0.051					
4^1A_u	5.85	212	0.884					
1^3A_u	3.48	356		3.04	408	2.49	498	0.51
2^3A_u	4.63	268						
	Experimental							
	A_{max}[b] /eV	λ_{amax}[b] /nm	f[d]	T_0[b] /eV	λ_0[b] /nm	E_{max}[b] /eV	λ_{emax}[b] /eV	τ_{rad} / s
1^1A_u	3.77	329	~0.4	~3.54	~350	3.05	407	1.2×10^{-8}
2^1A_u	4.56	272	~0.2					
4^1A_u	5.76	215	~1.0					

[a]Computed vertical absorption energy (E_{VA}) and wavelength (λ_{VA}), electronic band origin energy (T_e) and wavelength (λ_e), vertical emission energy (E_{VE}) and wavelength (λ_{VE}), oscillator strength (f), radiative lifetime (τ_{rad}). [b]Experimental absorption band maxima energy (A_{max}) and wavelength (λ_{amax}), band origin energy (T_0) and wavelength (λ_0), emission band maxima energy (E_{max}) and wavelength (λ_{emax}) in hexane. [c]Maximum of the transient absorption experiments: 2.42 eV (512 nm). At the 1^1A_u optimized geometry the most intense transition from 1^1A_u is computed to the 7^1A_g state at 2.69 eV with oscillator strength 0.352.[d]Normalized using molar absorption coefficients.

wave function: H-1 → L (15%) and H → L+1 (70%). A further two allowed transitions have been obtained at 186 (6.65) and 183 nm (6.76 eV), the highest with a large oscillator strength of 0.746 and described by the H → L+1 (48%) configuration. Additionally, nine states of both ^3A_g and ^3A_u symmetry have been computed in the energy range $356 - 215$ nm ($3.48 - 5.76$ eV). The two lowest singlet-triplet transitions have been obtained toward the 1^3A_u and 2^3A_u states at 356 (3.48) and 268 nm (4.63 eV), described by the H → L and H → L+1 configurations, respectively. Only one triplet state therefore lies below the lowest singlet-singlet transition. The spin-orbit coupling (SOC) terms obtained between singlet and triplet states were, in general, large (from 3 to 5 cm^{-1}) for states of the same spatial symmetry. Considering, however, the low-energy region of the spectrum where only the 1^1A_u state becomes largely populated, the closest-lying triplet states lie more than $0.5 - 0.7$ eV away in energy, making the energy gap too large to expect an efficient intersystem

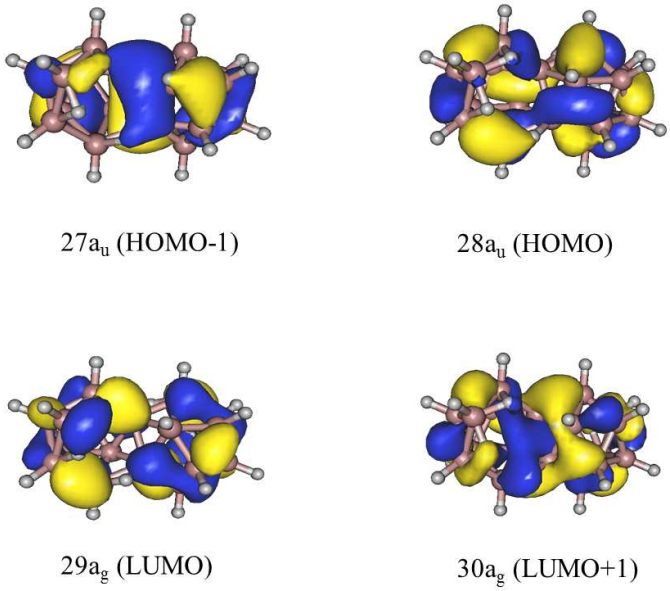


Figure 6.10: Most relevant CASSCF molecular orbitals for the *anti*-$B_{18}H_{22}$ isomer.

crossing process at the ground-state Franck-Condon (FC) region. This process can be, however, enhanced in other areas of the PEH.

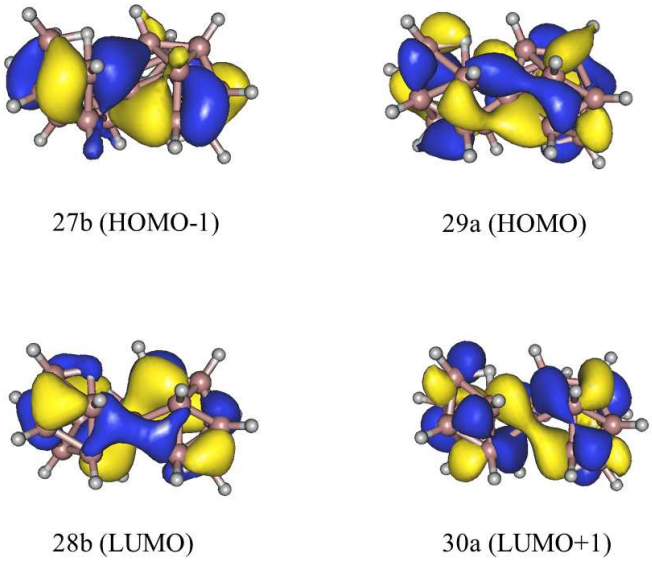


Figure 6.11: Most relevant CASSCF molecular orbitals for the *syn*-$B_{18}H_{22}$ isomer.

Table 6.14 compile the CASPT2 results and experimental data related to the absorption spectrum of the *syn*-$B_{18}H_{22}$ system, which displays a C_2 point group symmetry for its ground (S_0) state (1^1A) minimum at the B3LYP level of calculation. The *syn* isomer was found to be 0.42 eV (9.7 kcal·mol^{-1}) less stable than the *anti* counterpart at the CASPT2 level of calculation. The two medium-intensity bands observed in the UV absorption spectrum in hexane at 308 nm (4.03 eV) and 229 nm (5.41 eV) can be assigned to the lowest-lying transition to the 1^1B and 3^1B states computed at 291 (4.26 eV) and 215 nm (5.76 eV) with oscillator strengths 0.338 and 0.251, respectively. According to their CASSCF wave function, these two states can be made to correspond to Platt's ^1L_a (H $\rightarrow$ L 85%) and ^1B_a (H-1 $\rightarrow$ L+1 74%) states. With regard to Platt's b-type states, in this system the mixture takes place for the 3^1A and 4^1A states between the H $\rightarrow$ L+2 and H-2 $\rightarrow$ L configurations. Figure

Table 6.14: Theoretical (CASPT2) absorption and emission spectra of *syn*-$B_{18}H_{22}$.

	CASPT2							
State	E_{VA}[a] /eV	λ_{VA}[a] /nm	f[a]	T_e[a] /eV	λ_e[a] /nm	E_{VE}[a] /eV	λ_{VE}[a] /nm	τ_{rad} / s
1^1B	4.26	291	0.338	4.03[c]	308[c]	3.36[c]	369[c]	
4^1A	5.56	223	0.158					
3^1B	5.76	215	0.251					
1^3B	3.76	329		3.31	375	2.67	464	0.96
1^3A	4.13	300						
	Experimental							
	A_{max}[b] /eV	λ_{amax}[b] /nm	f[d]	T_0[b] /eV	λ_0[b] /nm	E_{maz}[b] /eV	λ_{emax}[b] /eV	τ_{rad} / s
1^1B	4.03	308	~0.6					
4^1A,3^1B	5.41	229	~0.5					

[a]Computed vertical absorption energy (E_{VA}) and wavelength (λ_{VA}), electronic band origin energy (T_e) and wavelength (λ_e), vertical emission energy (E_{VE}) and wavelength (λ_{VE}), oscillator strength (f), radiative lifetime (τ_{rad}). [b]Experimental absorption band maxima energy (A_{max}) and wavelength (λ_{amax}) in hexane. [c]The minimum was found along the relaxation path of the 1^1B state only by restricting the symmetry to C_2. Otherwise the optimization leads directly to the crossing with the ground state $(S_0/S_1)_{CI}$. No fluorescence emission is therefore expected.[d]Normalized using molar absorption coefficients.

6.11 displays the corresponding CASSCF molecular orbitals. Other transitions will contribute to a lesser extent to the bands intensity, such as the excitation to the 4^1A state at 223 nm (5.56 eV), which also displays a large oscillator strength and presumably participates in the 229 nm band. In the experimental spectrum (Fig. 6.5) there is an intense band peaking at energies higher than 200 nm (6.2 eV), which corresponds to the series of intense transitions computed in the 190 nm (6.5 eV) region. In contrast to the apolar *anti*-$B_{18}H_{22}$ isomer, the C_2 symmetry *syn*-$B_{18}H_{22}$ species has a polar character. The dipole moment of the ground state is computed to be 2.45 D, higher than for any of the singlet states computed below 6 eV. Corresponding blue-shifts can be therefore expected for the observed absorption bands, although the lowest-lying feature to the 1^1B state (having a dipole moment of 2.12 D) will not be strongly affected. A number of triplet excited states have been also computed for *syn*-$B_{18}H_{22}$. In this case, two triplet-transitions lie below the lowest-lying singlet state at 291 nm (4.26 eV): the $1^1\text{A} \rightarrow 1^3\text{B}$ transition at 329 nm (3.76 eV) and the $1^1\text{A} \rightarrow 1^3\text{A}$ at 300 nm (4.13 eV). In general, and except for the interaction between the ground singlet and the triplet states, the singlet-triplet SOC elements are smaller for the *syn*-$B_{18}H_{22}$ than for the *anti*-$B_{18}H_{22}$ isomer. In the FC region, however, there is a small gap between the most populated state of the low-energy tail of the spectrum, 1^1B, and the 1^3A state (0.13 eV). These states display a SOC term of 1.2 cm^{-1}, which, combined with the small energy gap, indicate a potential channel for ISC transfer toward the triplet manifold in this area of the PEH.

Fluorescence

Tables 6.13 and 6.14 compile the most important computed (CASPT2) and experimental data describing the photophysics of the two *anti* and *syn* isomers of $B_{18}H_{22}$, including absorption, emission, and decay properties. Figure 6.12 summarizes the proposed photophysics based on the experimental and computed results. Regarding the *anti*-$B_{18}H_{22}$ isomer, one excited state of singlet nature and another of triplet character seem to be the main protagonists of

its photophysics in the low-energy region of the spectrum. After being populated, the system in the 1^1A_u (S_1) state is expected to evolve toward the state minimum, computed adiabatically (T_e, electronic minimum of S_0 to electronic minimum of S_1) at 341 nm (3.63 eV). This energy corresponds to the band origin, both for the absorption and emission band (except for the solvent effects) measured around 364 nm (3.41 eV) in hexane (Fig. 6.5, Table 6.13). The emission energy computed from the S_1 state minimum vertically to the S_0 state at 426 nm (2.91 eV) nicely agrees to the fluorescence emission maximum observed at 407 nm (3.05 eV). As *anti*-$B_{18}H_{22}$ yields an intense emission, the Strickler-Berg approximation [169] can be safely applied to compute the radiative lifetime of $\tau_{rad} = 5.6 \times 10^{-9}$ s, a time short enough to correspond to an intensely absorbing state such as 1^1A_u. The experimental τ_{rad} can be calculated using the expression τ_F/Φ_F, to give 1.2×10^{-8} s, a value that is not very far from the calculated radiative lifetime.

The presence of a minimum along the main decay pathway of the S_1 state is an initial indication of the intrinsic emissive character of *anti*-$B_{18}H_{22}$ and the relatively long lifetime of the state. Indeed, time-resolved fluorescence and femtosecond transient absorption experiments recorded in hexane at 512 nm (2.42 eV) demonstrate a slow 11.2 ns decay for *anti*-$B_{18}H_{22}$, whereas the *syn* species has a corresponding absorption at 475 nm (2.61 eV), and a comparatively much faster S_1 decay of 75 ps (Fig. 6.6, 6.7). The short lifetime in the latter case is in agreement with the lack of fluorescence emission, pointing to fast internal conversion toward the ground state within *syn*-$B_{18}H_{22}$. The differing photophysics of the two isomers can be directly rationalized through the PEH decay profiles along the spectroscopic S_1 state (Fig. 6.12). Modern photochemistry relates fast or ultrafast internal conversions (IC) to the presence of accessible conical intersections (CI) between the corresponding excited states, implying both electronic energy degeneracy and strong nonadiabatic coupling. In the present situation, the strong fluorescence emission and a relative long lifetime points, in *anti*-$B_{18}H_{22}$, to a deep S_1 minimum and a CI between S_0 and S_1, $(S_0/S_1)_{CI}$, that cannot be easily accessed with the initially available excess energy (315 nm, 3.93 eV). The present CASPT2//CASSCF calculations have computed the lowest-lying $(S_0/S_1)_{CI}$ energy crossing adiabatically

at 299 nm (4.14 eV), representing an energy barrier of 0.51 eV with respect to the S_1 minimum. The expected slopped structure of the CI is indicative of the decay properties of the *anti* isomer toward the ground state, which is slower and less probable than that in the *syn* species, both because the CI cannot be efficiently reached and because the wave packet is expected to undergo many crossings and recrossings prior to decay toward the ground state, thus explaining the major radiative character of the deactivation and its nanosecond lifetime. Conversely, the *syn* isomer lacks measurable fluorescence emission and displays a much faster picosecond decay. Optimization of the S_1 state of *syn*-$B_{18}H_{22}$ leads to a minimum if the molecular symmetry is restricted to the C_2 point group, and directly to the crossing with the ground state once the system is allowed to break its symmetry. The PEH region close to the mini-

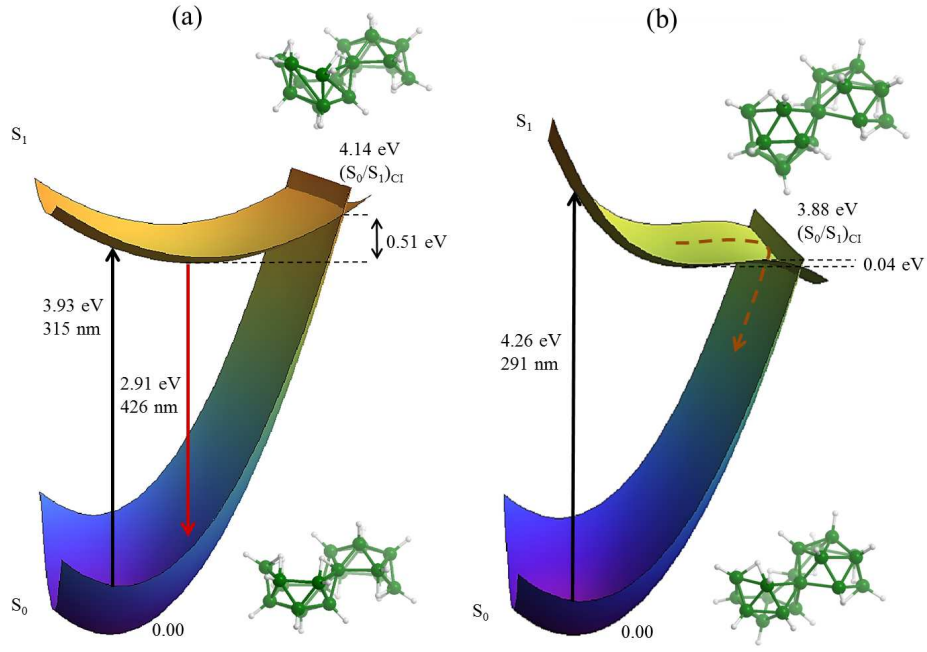


Figure 6.12: Photophysics, potential energy hypersurfaces, and structures of the ground state (S_0) and lowest excited singlet state (S_1) for *anti*-$B_{18}H_{22}$ (a) and *syn*-$B_{18}H_{22}$ (b) based on experimental and CASPT2 results.

mum and close to the crossing is extremely flat. The calculation of the MECP leads to a $(S_0/S_1)_{CI}$ crossing lying adiabatically at 320 nm (3.88 eV) from the ground state minimum. With the consideration that, for the *syn* isomer, the excitation energy to the 1^1B (S_1) state at the FC region is 291 nm (4.26 eV), the system initially has enough excess energy to rapidly and efficiently reach the CI and to decay to the ground state in a nonradiative way, thus explaining the absence of fluorescence emission and the picosecond character of the deactivation pattern observed for the *syn* isomer. A Linear Interpolation in Internal Coordinates (LIIC) was performed between these two structures leading to a very small barrier of 1 kcal·mol^{-1}. In practice this means that no effective minimum can efficiently trap the energy along the main deactivation path, pointing to a barrierless profile towards the CI.

Triplet States

With regard to the location of the triplet states in *anti*-$B_{18}H_{22}$, the lowest-energy triplet T_1 state lies 0.45 eV below S_1 at the FC region, showing a similar relaxation profile than the latter, is placed at the S_1 state minimum at 0.50 eV below the singlet state. Singlet state minima are structures favorable for intersystem crossing (ISC) to take place because the energy is trapped for a long enough time to allow ISC processes to occur in competition to generally faster internal conversion (IC) processes. In this case, however, all the triplet states seem, in terms of their respective energies, far apart from S_1. As the system approaches the $(S_0/S_1)_{CI}$ structure, however, the S_1-T_1 energy gap decreases, enhancing the probability for ISC to take place. Although *anti*-$B_{18}H_{22}$ has a high fluorescence quantum yield of 0.97, the additional population of the triplet manifold seem confirmed by the observed ability to produce $O_2(^1\Delta_g)$ in the presence of molecular oxygen. Production of $O_2(^1\Delta_g)$ usually proceeds through photosensitization from a triplet state of a donor system via an energy transfer process [260]. The more efficiently the triplet state is populated and the larger the coupling is with the resulting singlet oxygen state, then the higher will be the $O_2(^1\Delta_g)$ production [261]. The state of the donor res-

ponsible for the $O_2(^1\Delta_g)$ formation, presumably the T_1 triplet state, has a lifetime of ~ 6 μs affording a quantum yield Φ_Δ below 0.01 in hexane solutions. The calculated adiabatic energy for T_1 of the *anti*-$B_{18}H_{22}$ donor is 408 nm (3.04 eV), more than 2 eV higher than the 0.97 eV required to populate the $O_2(^1\Delta_g)$ state through a triplet-triplet energy transfer (TET) mechanism [176, 260]. In these cases, when the triplet donor lies higher in energy than the final acceptor state, the TET process is exothermic and diffusion controlled, with molecular oxygen behaving as a rigid, classical acceptor [262]. Using the computed radiative lifetime for T_1 and low Φ_Δ (< 0.01) as a measure of the donor triplet state formation yield, the expected lifetime of the T_1 state of *anti*-$B_{18}H_{22}$ rises to $\sim 50 - 500$ μs under oxygen-free conditions.

With regard to *syn*-$B_{18}H_{22}$, the 1^1B (S_1) and 1^3A (T_2) states remain close-lying both at the FC structure and the constrained C_2 S_1 minimum, opening the possibility of a favorable ISC, especially as the S_1 minimum is not available along the main relaxation path. Considering that the S_1 PEH is somewhat flat near the CI, contributions from an ISC process toward T_1 is not completely unlikely, in particular because the SOC element between S_1 and T_1 is 3.1 cm^{-1}. The evidence of the $O_2(^1\Delta_g)$ formation suggests the involvement of the T_1 triplet states.

Geometrical changes at stationary points of the PEH

Figure 6.13 shows the structure and numbering of (a) *anti*-$B_{18}H_{22}$ (C_i symmetry) and (b) *syn*-$B_{18}H_{22}$ (C_2 symmetry) as proposed by Lipscomb [40, 42] and Todd [263]. The boron atoms related by symmetry operations in these clusters are signified by number/'primed' number pairs. A study of the B-B distance changes in the different stationary points of the PEH has been carried out for *anti*-$B_{18}H_{22}$ and *syn*-$B_{18}H_{22}$ as the geometry evolves from the S_0 ground state to the stationary points T_1, S_1 and $(S_0/S_1)_{CI}$, respectively.

The geometrical change profiles for $S_0 \rightarrow T_1/S_1$ in *anti*-$B_{18}H_{22}$ and *syn*-$B_{18}H_{22}$ are very similar as explained below. It is defined Δd as the difference between B-B distances in the geometrical change A $\rightarrow$ B. For *anti*-$B_{18}H_{22}$, the $S_0 \rightarrow T_1/S_1$ changes follow very similar profiles, with a maximum B-B contraction of $\Delta d \sim 0.15/0.13$ Å (T_1/S_1) for B_7-B_8 and the corresponding symmetry-equivalent $B_6-B_{8'}$ (see Fig. 6.13a) followed by values of Δd below 0.10 Å. As for B-B elongations, the maximum value corresponds to B_8-B_9 (and equivalent $B_{8'}-B_{9'}$) with $\Delta d \sim 0.08$ Å.

In the case of *syn*-$B_{18}H_{22}$, the geometrical change profile for $S_0 \rightarrow T_1/S_1$ follows a similar behavior to that of the *anti*-$B_{18}H_{22}$ isomer. The maximum value for B-B shrinkage is in B_7-B_8 ($B_7-B_{8'}$) with $\Delta d \sim 0.12/0.09$ Å (T_1/S_1), the same boron atoms involved as for the *anti*-$B_{18}H_{22}$ isomer. The maximum B-B elongation corresponds to B_5-B_6 with $\Delta d \sim 0.10/0.11$ Å (T_1/S_1), followed by elongations smaller than $\Delta d \sim 0.03$ Å.

The most significant changes in the molecular geometries of both *anti*- and *syn*-$B_{18}H_{22}$ occur at the conical intersection points along their PEH. At these positions the calculated geometries of both isomers are distorted and completely lose their ground state symmetries. Diagrams (c) and (d) in Fig. 6.13 illustrate the most significant geometrical changes: Wavy lines between boron atoms indicate B-B connectivity elongations that extend to clearly non-bonding distances, and heavy lines indicate significant B-B connectivity contractions. Overall, the majority of the B-B connectivity distances for *anti*- and *syn*-$B_{18}H_{22}$ molecules at their $(S_0/S_1)_{CI}$ points remain essentially unchanged from those distances at their S_0 ground states. Significantly, however, the changes that do occur are large and seem to be concentrated in a small number of molecular vibrational modes. In the case of *anti*-$B_{18}H_{22}$ (see Fig. 6.13, diagram (c)) the most significant geometrical changes are B-B connectivity elongations for $B_{1'}-B_{10'}$ ($\Delta d \sim 0.90$ Å), $B_{2'}-B_{5'}$ ($\Delta d \sim 0.70$ Å) and $B_7-B_{5'}$ ($\Delta d \sim 0.55$ Å), and a B-B contraction along the $B_{5'}-B_{10'}$ vector of approximately 0.25 Å. The resulting deformation of the $B_7-B_{2'}-B_{5'}$ structural deltahedron leads to disruption of the $H_{7,5'}$ bridging hydrogen (3-center, 2-electron) bond and the establishment of a B-H bond of clear terminal (2-

center, 2-electron) character at the B_7 position. The case for *syn*-$B_{18}H_{22}$ (see Fig. 6.13, diagram (d)) is considerably simpler. Here there is a similar contraction of the $B_{5'}-B_{10'}$ connectivity of about 0.26 Å that seemingly disrupts the $H_{9',10'}$ bond to the bridging hydrogen, leaving an *endo*-like 2-center, 2-

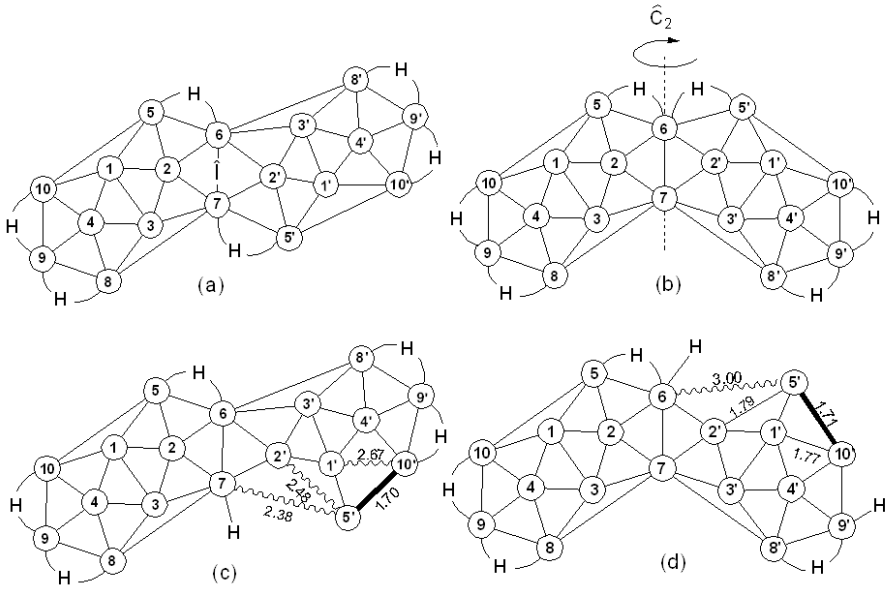


Figure 6.13: Atom labels for the systems: (a) *anti*-$B_{18}H_{22}$ molecular geometry at S_0 ground state, C_i symmetry. Only boron labels are shown together with 'bridging' and '*endo*' hydrogen atoms. Terminal hydrogen atoms are omitted for clarity. The inversion operator $\hat{\mathbf{i}}$ (shown in the center of the molecule) relates labels A with $A' \rightarrow \hat{\mathbf{i}}(A) = A'$, except for B_6 and B_7, where $\hat{\mathbf{i}}(B_{6,7}) = B_{7,6}$. (b) *syn*-$B_{18}H_{22}$ molecular geometry at S_0 ground state, C_2 symmetry. The $\hat{\mathbf{C}}_\mathbf{2}$ axis of rotation relates labels A with $A' \rightarrow \hat{\mathbf{C}}_\mathbf{2}(A) = A'$, except for B_6 and B_7, where $\hat{\mathbf{C}}_\mathbf{2}(A) = A$. (c) *anti*-$B_{18}H_{22}$ molecular geometry at $(S_0/S_1)_{CI}$. (d) *syn*-$B_{18}H_{22}$ molecular geometry at $(S_0/S_1)_{CI}$. Significantly elongated B-B connectivities showed as a wavy line, and significantly contracted B-B connectivities as a heavy line, both with respective connectivity distances written in units of Ångström to two decimal places.

electron $B_{9'}-H_{9'}$ bond in its stead. However, most strikingly of all, the S_0 $\rightarrow$ $(S_0/S_1)_{CI}$ progression for *syn*-$B_{18}H_{22}$ results in only one significant B-B elongation; an approximate increase of 1.2 Å from 1.8 to 3.0 Å in the $B_6-B_{5'}$ connectivity.

Thus, the distinguishing factor when comparing the geometries of the conical intersections between *anti*-$B_{18}H_{22}$ and *syn*-$B_{18}H_{22}$, is the presence of a conspicuous B-B elongation well above 1 Å as compared to the S_0 ground state geometry for *syn*-$B_{18}H_{22}$. The *anti*-$B_{18}H_{22}$ isomer also undergoes significant elongations for $(S_0/S_1)_{CI}$, but, in this case, several B-B connectivities are below 1 Å, such that the absorbed energy in S_1 is redistributed vibrationally into several regions of the molecule rather than (almost completely) into a single vibrational mode, as in *syn*-$B_{18}H_{22}$, allowing for a minimum in the S_1 PEH (Fig. 6.12a). From the S_1 minimum the *anti* isomer undergoes a radiative decay to its ground state via fluorescence, a feature not found for the *syn* isomer.

6.4. Proton/Hydrogen transfer processes in the Adenine-Thymine and Guanine-Cytosine base pairs

DNA encodes the genetic information responsible for the development and functioning of living organisms. Since the discovery of the chemical composition of DNA and the associated structural arrangement in a double-helix form [47, 264], the profound understanding of its properties has constituted an intriguing challenge to the scientific community. Especial attention has been focused on DNA damage caused by UV radiation, and within this context, theoretical and experimental studies have been actively performed [265, 266, 267].

Effects of DNA damage can be severe because it may interrupt replication or originate transitions or transversions mutations, altering so the ordering of the nucleic acid bases (NABs), and therefore modifying the genetic code [268, 269]. The tautomeric forms in the guanine-cytosine (GC) base pairs have been recently suggested to be responsible for the universal mutation guanine-cytosine to adenine-thymine, GC $\rightarrow$ AT, frequently found in bacteria, fungi, plants, and animals [270]. In order to deal with DNA lesions, nature has developed enzymatic mechanisms, which are involved in the repair process. The interplay between both mutagenic and DNA-repair processes constitutes the basic framework in which the evolution of life is feasible.

In the last decade, different research groups have focused on computational studies aiming at understanding the significance of the chemical physical properties of the natural selected canonical NABs, which are more stable than their tautomeric forms under UV irradiation, and on underlying mechanisms preserving the genetic code and avoiding mutations [271]. To avoid undesired mutations, nature has developed nucleobases with very short excited-state lifetimes, which is known to be an intrinsic property of the isolated canonical nucleobases and is related to nonadiabatic processes that, ultimately, leads to the relaxation of the $^1\pi\pi^*$ excited states to the ground state, S_0, in a barrierless pathway [272, 273, 274, 275, 276]. Due to such extremely fast relaxation

processes, usually in the femtosecond timescale, the excited species does not exist enough time to undergo chemical reactions. Additionally, femtosecond pump-probe experiments with the canonical NABs reported multiexponential decay channels in the femtosecond and picosecond timescales, a strong evidence for the stability of the natural NABs [271, 277, 278, 279].

New photochemical paths have been more recently determined in π-stacked NABs [280, 281], which compete with the photostable routes present in the nucleobase monomers. The proximity of the aromatic molecules may facilitate the formation of excimers (excited dimers). The most striking feature of these excimers is the rising of long-lived emissive features not existent in the isolated nucleobases [271]. Due to their long lifetimes, they are supposed to be the key for intrastrand photolesions [47, 281, 282, 283], which are the result of cycloaddition photoreactions in DNA molecules, generating photodimers and photoadducts [266, 284, 285]. Among the photodimers (cyclobutane pyrimidine dimers, CPDs), the most common lesion involving a single DNA strand corresponds to the dimerization of two stacked thymine molecules. On the other hand, the cycloaddition of two cytosine molecules, known as 6-4 cytosine adducts, is the photoadduct of major impact.

Despite relevant information can be obtained from the studies with the isolated NAB monomers and intrastrand π-stacked dimers, further aspects can be understood by exploring the photochemistry of the DNA base pairs. In 1953, Watson and Crick found the B-DNA structure as the most common arrangement of the genetic code [47] and proposed that the structure of DNA is composed by two nucleotide chains organized in double-helix manner around an axis, stabilized by hydrogen bonds formed between the canonical purine and pyrimidine bases: adenine-thymine (AT) and Guanine-Cytosine (GC) (see Fig. 6.14).

Early in 1963, Löwdin introduced the hypothesis that the double proton or hydrogen transfer reaction in DNA base pairs could be a possible source of spontaneous mutations [54, 55]. Since this work, the double proton/hydrogen

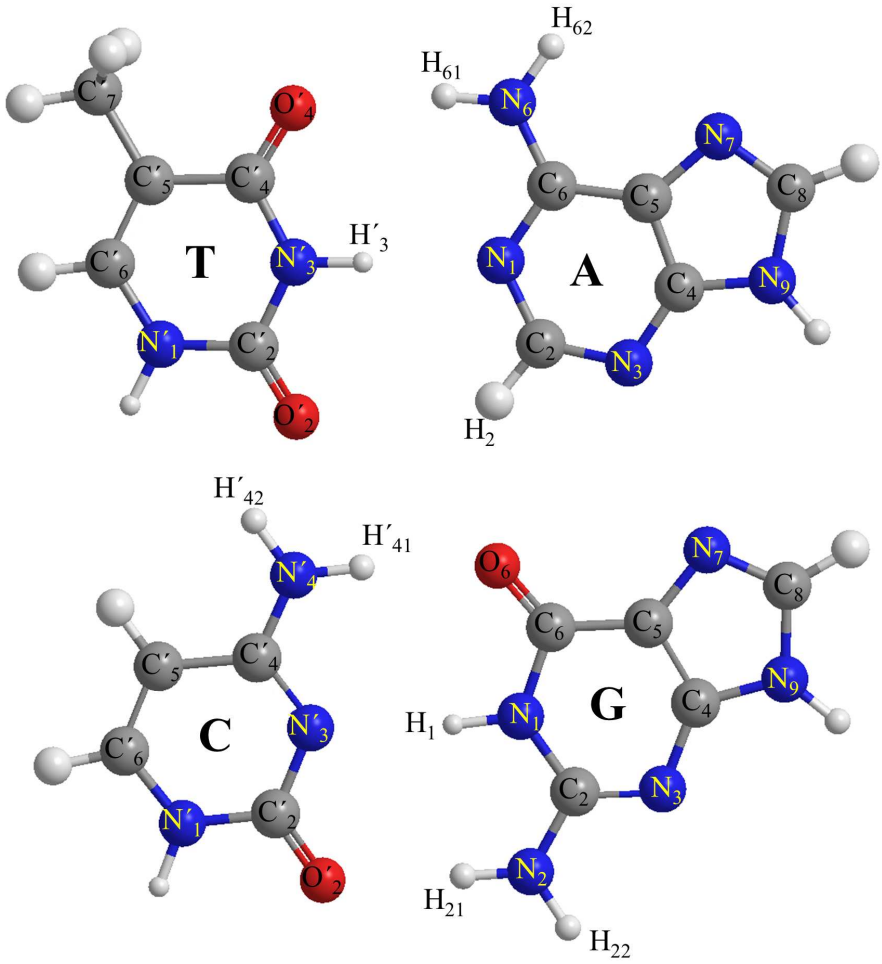


Figure 6.14: Atom labeling and numbering for the Adenine-Thymine and Guanine-Cytosine base pairs.

transfer induced by photoexcitation in model DNA base pairs has been studied both experimental [66, 67, 69, 70, 72, 74, 85] and theoretically [75, 76, 77, 79]. In order to investigate whether the reaction mechanism is stepwise or concerted, and the nature of the intermediate species, Kwon et al. employed femtosecond dynamic techniques in the study of the 7-azaindole [74], con-

cluding that the stepwise pathway studied by early theoretical works [79] is favored.

Double transfer of a proton/hydrogen has also been considered as a possible mechanism to explain the photostability of the WC DNA base pairs [74, 80, 81, 82, 83, 286]. At first, the proton H_{61} in AT or H_1 in GC migrates from the purine to the pyrimidine monomer (see Fig. 6.14). Next, the proton returns to the purine (A or G) moiety, which supports the stepwise mechanism. Domcke and co-workers have proposed that the proton transfer takes place after population of a charge transfer (CT) state in the canonical base pairs [80, 81, 82, 83]. The radiationless deactivation could be mediated by a conical intersection (CI) between the charge transfer state and the ground state, $(S_0/S_1)_{CI}$, taking place while the proton is transferred to the pyrimidine moiety. The charge separation in the ground-state potential energy surface at the $(S_0/S_1)_{CI}$ crossing allows the back transfer of the second proton to the purine moiety. The mechanism is known as electron-driven double transfer of a single proton.

As mentioned in section 1.4, despite the fact that the studies on models of the DNA base pairs clearly pointed out to an active participation of the double proton/hydrogen transfer in the photochemistry of the systems, the relevance of base pairing in the excited-state dynamics of the DNA molecule is still uncertain [86, 88, 89, 90, 91]. Experiments in alternating DNA duplexes poly(dGdC)·poly(dGdC) using upconversion spectroscopy found shortened fluorescence lifetimes as compared to those measured in the mononucleotides [86]. Nevertheless, transient absorption spectra recorded for $d(GC)_9 \cdot d(GC)_9$ double strand determined that the ground-state recovery process is longer than in the mononucleotides, suggesting that the decay occurs through the formation of exciplexes [88]. Additionally, a combined experimental and theoretical study on the GC base pair in solution with chloroform concluded that the proton/hydrogen transfer is a nonefficient process for energy decay.

A model based on the characterization of the energy-decay channels involving proton/hydrogen transfer in the AT and GC base pairs seems timely in order to establish the operative mechanisms for photostability and tautomerism. In the present thesis, the proton/hydrogen transfer of the isolated AT and GC base pairs in the excited state have been studied by mapping the two-dimensional potential energy surfaces along the hydrogen-bond reaction coordinates, evaluating different possible competitive deactivation pathways. Additionally, the main equilibrium structures for the double proton/hydrogen phenomena in the isolated GC base pair are also characterized in a DNA double helix of 18 alternating GC base pairs in water (see Fig. 6.15), which allow comparisons with previous experiments on DNA duplexes with the same nucleotide sequence. Insights into the dynamics of the GC base pair *in vaccuo* are also addressed. Previous works on the dynamics of the system have been reported, although at levels of theory not able to describe the overall photochemistry of the GC base pair [89].

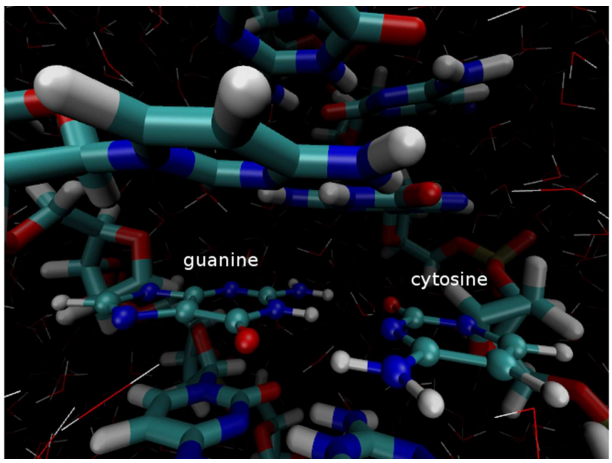


Figure 6.15: Perspective view of the GC base pair in a hydrated DNA double helix of 18 alternating base pairs, namely $d(GC)_9 \cdot d(GC)_9$.

The strategies designed to explore the energy-decay routes of the AT and GC base pairs are shown below together with the methods employed to characterize the electronic structure and the computational techniques used in the semiclassical dynamics calculations.

The energy-decay paths of the AT base pair

Ground as well as excited states of LE (locally excited) and CT character have been described. Both stepwise and concerted pathways have been considered in a schematic manner as shown in Fig. 6.16. In short, the canonical base pair, hereafter WC AT, at the ground-state equilibrium geometry absorbs UV radiation. In the Franck-Condon (FC) region, the energy is transferred to the low-lying $\pi\pi^*$ bright state. The T-based keto-enol tautomer (hereafter, $\mathrm{TAU}_{(\mathrm{AT})}$) of the AT base pair can be formed when a double proton/hydrogen transfer at the $\mathrm{N}_3'\mathrm{H}_3' \cdots \mathrm{N}_1$ and $\mathrm{O}_4' \cdots \mathrm{H}_{61}\mathrm{N}_6$ sites results in the 1H-9H-6-Iminopurine/2-oxo-4-hydroxypyrimidine. In the concerted or symmetric pathway, the H atoms are transferred synchronously ($\mathrm{CONC}_{(\mathrm{AT})}$), connecting the canonical AT base pair and the tautomeric form as shown in Fig. 6.16. Meanwhile, two asymmetric pathways, via intermediates of different nature (either neutral or ionic), can be relevant to the photochemistry of the dimer. The neutral species ($\mathrm{INT1NEU}_{(\mathrm{AT})}$) feature a CT electronic structure, whereas the ionic ones ($\mathrm{INT1ION}_{(\mathrm{AT})}$ and $\mathrm{INT2ION}_{(\mathrm{AT})}$) correspond to LE states.

The energy-decay paths of the GC base pair

The ground plus LE and CT excited states have been characterized in the GC base pair. Both stepwise and concerted pathways have been studied in such a way as shown in Fig. 6.17 - 6.19. As occurs in AT, the canonical GC base pair, hereafter WC GC, at the ground-state equilibrium geometry is irradiated by UV radiation and, in the FC region, the energy is transferred to the low-lying $\pi\pi^*$ bright state. Two tautomeric configurations of the GC base pair can be potentially formed from this point depending on the hydrogen bonds involved in the proton/hydrogen transfer processes. Whereas a double proton/hydrogen transfer at the $\mathrm{N}_3' \cdots \mathrm{H}_1\mathrm{N}_1$ and $\mathrm{N}_4'\mathrm{H}_{41}' \cdots \mathrm{O}_6$ sites might result in the 9H-2-Amino-6-hydroxypurine/3H-2-oxo-4-Iminopyrimidine

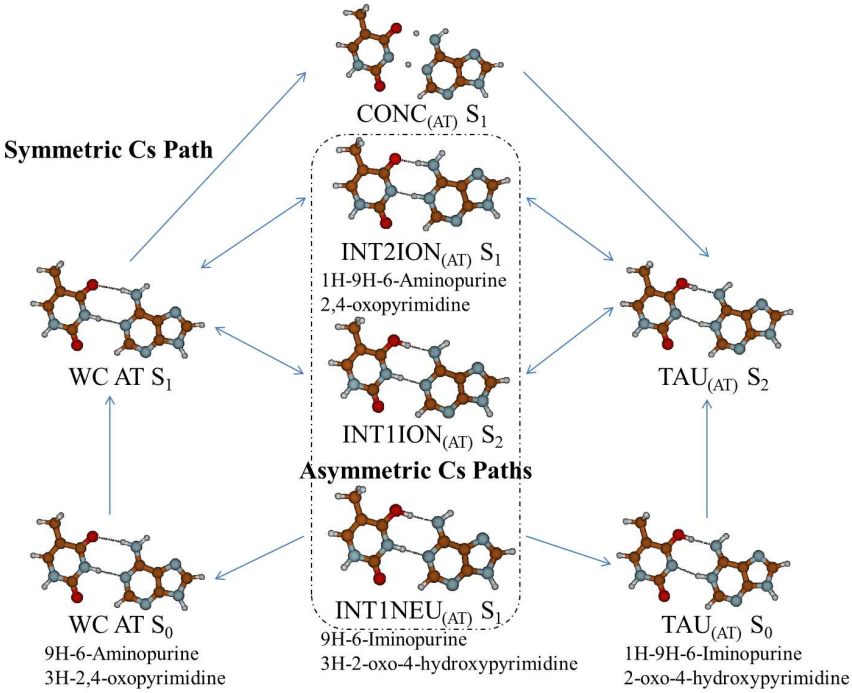


Figure 6.16: Scheme of the paths explored for the excited state double proton/hydrogen transfer process in the AT base pair. The atoms transferred are H_{61} of the amino group in A plus H'_3 in T. For clearness, the asymmetric C_s paths are enclosed in a dotted box.

or C-based imino-keto tautomers (hereafter, $TAU1_{(GC)}$), the 1H-9H-2-Imino-6-hydroxypurine/2-hydroxy-4-Iminopyrimidine or C-based imino-enol base-pair configuration (hereafter, $TAU2_{(GC)}$) can be produced after an analogous process at the $O'_2 \cdots H_{21}N_2$ and $N'_4H'_{41} \cdots O_6$ sites. In the concerted or symmetric route, the H atoms connect synchronously ($CONC1_{(GC)}$ and $CONC2_{(GC)}$) the canonical GC base pair and the respective tautomeric form as shown in Fig. 6.17 and 6.18. On the other hand, the asymmetric pathways, via either neutral or ionic intermediates, seem to have an important role in the photochemistry of the GC base pair. The neutral species ($INT1NEU_{(GC)}$ and $INT2NEU_{(GC)}$)

correspond to a CT electronic structure, whereas the ionic ones ($INT1ION_{(GC)}$ and $INT3ION_{(GC)}$) characterize LE states.

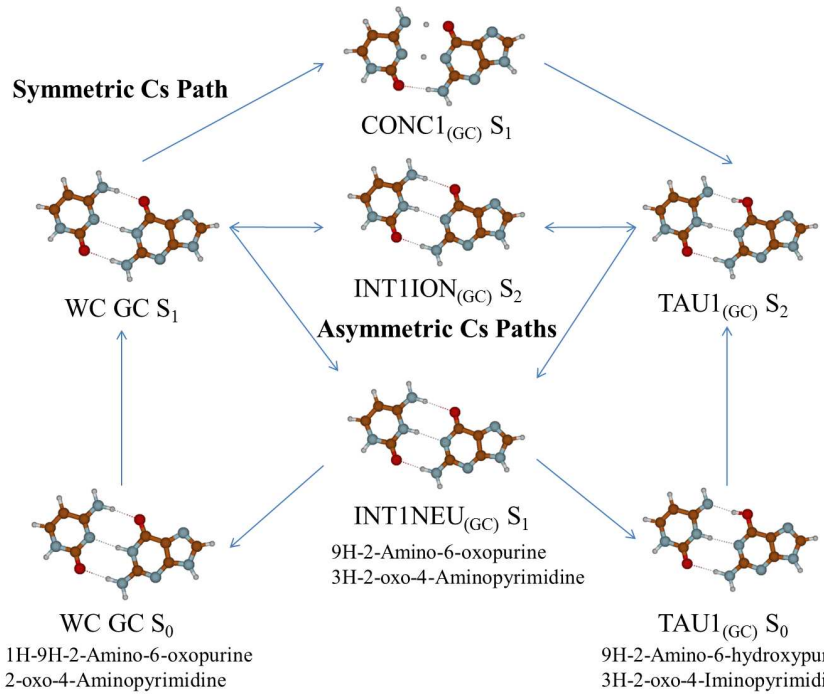


Figure 6.17: Scheme of the paths explored for the excited state double proton/hydrogen transfer process in the GC base pair. The atoms transferred are H_1 in G plus H'_{41} of the amino group in C.

Electronic Structure

Characterization of the lowest-lying $\pi\pi^*$ singlet excited states of the AT and GC base pairs has been performed with the complete active space self-consistent field second-order perturbation theory (CASPT2) method [9, 10], and the atomic natural orbital (ANO-S) [134] basis set of double-ζ quality plus polarization have been used, being described as C,N,O(10s6p3d)/H(7s3p) primitive gaussian functions contracted to C,N,O[3s2p1d]/H[2s1p]. A set of restricted

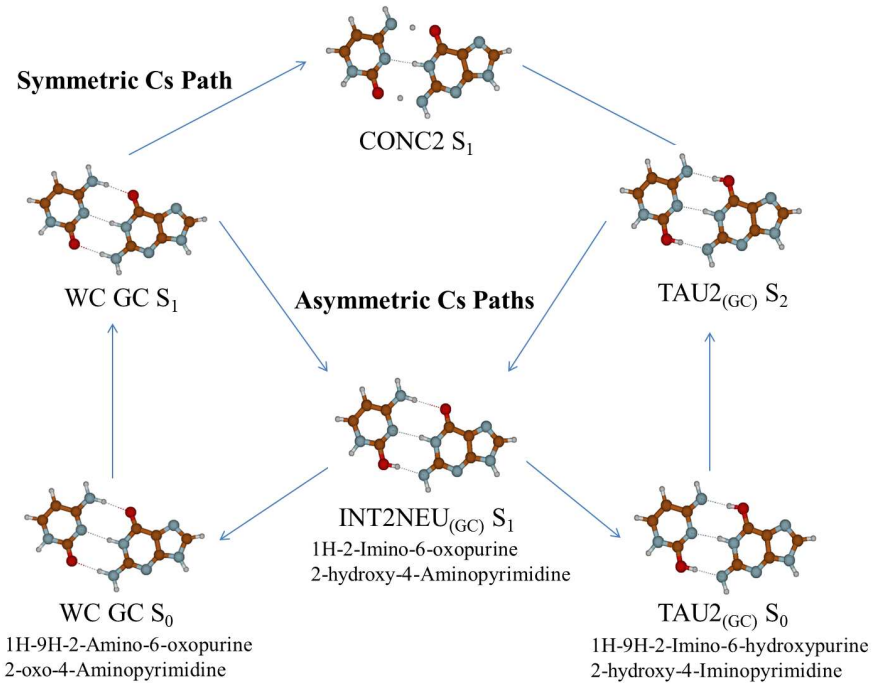


Figure 6.18: Scheme of the paths explored for the excited state double proton/hydrogen transfer process in the GC base pair. The atoms transferred are H_{21} and H'_{41} of the amino groups in both G and C, respectively.

active space second-order perturbation theory (RASPT2) calculations up to quadruple excited configurations constrained to C_s symmetry has been initially carried out in the canonical AT and GC base pairs and the tautomer $TAU1_{(GC)}$ in order to determine the main lowest-lying excited states that are involved in the photochemistry of the AT and GC base pairs. With the RAS2 subspace empty, the valence π and two lone pair n (RAS1) plus π^* (RAS3) molecular orbitals (MOs) have been employed in the calibration calculations. The lowest-energy $n\pi^*$ state for the WC AT base pair has been found at 4.90 eV and identified as an excited state localized on the thymine moiety. Nevertheless, two $\pi\pi^*$ LE states in adenine moiety lying lower in energy (4.73 and 4.78 eV) have a more relevant role in the photochemistry of the AT base pair.

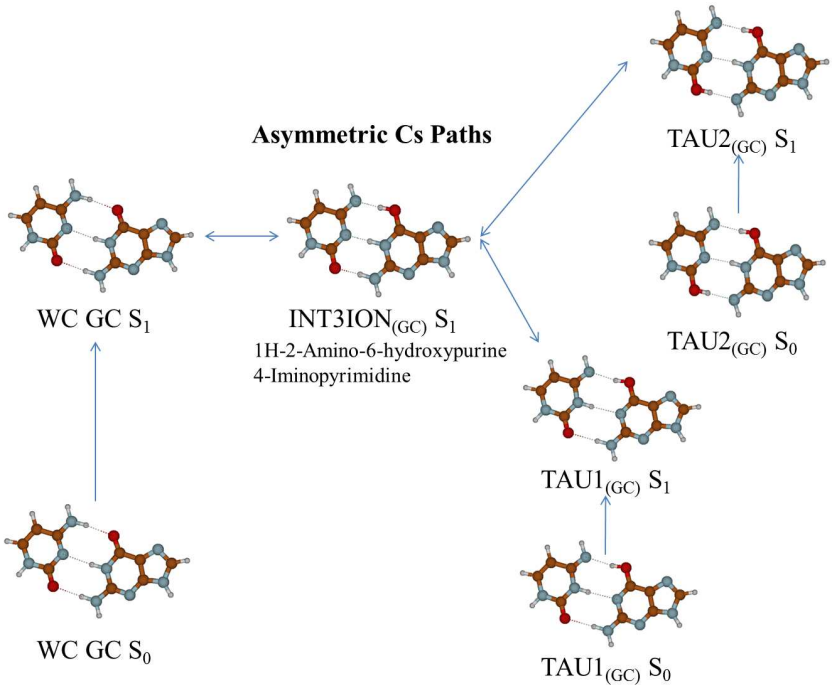


Figure 6.19: Scheme of the paths explored for the excited state double proton/hydrogen transfer process in the GC base pair. The atoms transferred are H'_{41} in the amino group of C plus either H_1 or H_{21} in G.

On the other hand, the highest-energy $\pi\pi^*$ states have been found at 5.20 eV and 5.53 eV for the canonical GC base pair and the tautomeric $TAU1_{(GC)}$ form, respectively, which are still below the energy of the lowest $n\pi^*$ states, placed at 5.91 eV and 5.90 eV. It consequently implies that in order to examine the photochemistry of the GC base pair just the $\pi\pi^*$ excited states have to be considered. Accordingly, the active space selected for the CASSCF/CASPT2 calculations comprises 12 π valence electrons distributed into 12 π active MOs [hereafter, CASSCF/CASPT2(12,12)], excluding the remaining occupied and unoccupied MOs with highest and lowest occupation numbers in the RASSCF computations, respectively. C_s symmetry constraint has been imposed in the CASSCF geometry optimizations and CASPT2 vertical energies. Oscillator

strengths have been computed from CASSCF dipole moments and CASPT2 energies. The conventional CASPT2 method has been employed (IPEA=0.0 au) [9]. In order to minimize weakly interacting intruder states, the imaginary level shift technique with a parameter of 0.2 au has been used. Two-dimensional potential energy surfaces (PESs) have been mapped along the reaction coordinates at the CASPT2 level by using linear interpolation in internal coordinates (LIIC) among the most relevant structures: canonical form, intermediates and tautomeric species.

The effect of the biological environment on the DNA base pairs have been evaluated for the GC base pair. The hybrid quantum mechanics/molecular mechanics (QM/MM) approach has been used to simulate properly the DNA surroundings. As QM core, the GC base pair (see Fig. 6.15) has been modeled by means of the *ab initio* CASSCF/CASPT2 method, using the same active space, basis set, and symmetry constraint as in the *in-vacuo* computation. The MM subsystem has been represented by using the AMBER99 force field [287, 288] in order to include the double helix strand, the aqueous media, and the external counterions. The double helix is composed by 18 alternating GC base pairs, $d(GC)_9 \cdot d(GC)_9$, and the DNA backbone of deoxyriboses and phosphates. The aqueous media is modeled by an octahedron box of 8618 water molecules surrounding the simulated DNA molecule, and the external counterions are represented by Na^+ monovalent cations. The initial structure employed in the QM/MM computations has been generated by using classical molecular dynamics (MD) simulations. The employed QM/MM approach takes into account the explicit electronic polarization of the base pair by the DNA double helix and the solvent. Geometry optimizations of the most important stationary points have been performed at the CASSCF/MM level for the GC base pair of interest, the related deoxyribose and phosphate groups, the adjacent nucleotides, and all the water molecules and Na^+ ions within a radius of 5 Å from the QM part. This allows for mutual geometric polarization between the QM moiety and its immediate surroundings. The CASPT2 energies of the lowest-energy states have been compared with the energy values obtained *in vacuo*.

Semiclassical Dynamics

Born-Oppenheimer dynamics calculations have been performed on the excited-state surface of the isolated GC base pair. Newton's equations have been solved by the velocity-Verlet algorithm [125] with a 1 fs time step. Analytical gradients have been computed on-the-fly at the CASSCF(12,12)/ANO-S C,N,O [$3s2p1d$]/H [$2s1p$] level of theory with the constraint of C_s symmetry. A group of 300 initial conditions have been generated according to the Wigner distribution for the quantum harmonic oscillator [289] and using the normal modes computed at the B3LYP/6-31G(d) level of theory. From this data, four representative sets of starting nuclear coordinates and momenta have been selected for subsequent dynamics simulations. The criteria for the selection of the sets are based on choosing the initial NH bond distances and scalar projections of the velocities of the H atoms along the inter-base-pair hydrogen bonds which potentially exhibited some of the dynamical aspects of the different relaxation mechanisms earlier established by the static models presented in this study.

The results of the present section are exposed in three parts. First, the mechanisms for proton/hydrogen transfer in the isolated AT and GC base pairs are explained. Next, the biological environment is taken into account in the study of the photochemistry of the GC base pair. And finally, some insights into the excited-state dynamics of the GC base pair are considered *in vacuo*.

6.4.1. Mechanisms of photostability and tautomers production in the isolated AT and GC base pairs

The equilibrium structures for the ground (S_0) and lowest-lying $^1\pi\pi^*$ excited states of the canonical (WC AT and WC GC), tautomeric ($TAU_{(AT)}$, $TAU1_{(GC)}$, and $TAU2_{(GC)}$), ionic intermediate ($INT1ION_{(AT)}$, $INT2ION_{(AT)}$,

$INT1ION_{(GC)}$, and $INT3ION_{(GC)}$), and neutral intermediate ($INT1NEU_{(AT)}$, $INT1NEU_{(GC)}$, and $INT2NEU_{(GC)}$) forms of the AT and GC base pairs have been obtained with the CASSCF method. The mapping of the PESs connecting these points at the CASPT2 level (Fig. 6.20 - 6.30) accounts for a qualitative picture of the competitive pathways for energy decay via proton/hydrogen transfer in the DNA base pairs. To determine the two-dimensional PESs, adiabatic curves are depicted. The excited states involve singlet LE and CT excitations and, within the former, both intra-purine (AA or GG) and intra-pyrimidine (TT or CC) electronic transitions are identified.

The photochemistry of the AT and GC base pairs start with the absorption of UV light. At the FC region of the canonical base pair, the brightest low-lying $^1\pi\pi^*$ state (oscillator strengths $f = 0.40$ and $f = 0.38$ for AT and GC, respectively) is a LE-state in the purine moiety with an excitation energy of 4.74 eV (see Fig. 6.20 - 6.30). From this point, the system will evolve toward the minimum of the S_1 PES. According to previous calculations on the isolated A and G bases, out-of-plane distortions in the six-member ring of A or G take place along the minimum energy path on the $^1\pi\pi^*$ excited state reaching a CI with the ground state that mediates the nonradiative energy decay in the monomer [272, 276]. In the DNA base pairs, however, the presence of the complementary NAB opens new photochemical paths for deactivating the excess energy after UV irradiation. Based on the present CASSCF/CASPT2 findings, three mechanisms are introduced at this point for the photochemistry of the DNA base pairs in the gas phase: a stepwise double proton transfer (SDPT), a stepwise double hydrogen transfer (SDHT), and a concerted double proton transfer (CDPT). Their relevance in the energy deactivation processes which follow the UV-irradiation of the dimer depends on the topology of the PESs along the reaction paths.

Stepwise double proton-transfer (SDPT) mechanism

The LE-state in the purine moiety is proposed as the driver of the photochemistry in this mechanism. Due to the local excitation, the species exchanged between the NABs are protons (not hydrogens). Five cases of double proton transfer have been studied under these conditions, which involve different hydrogen bonds, as depicted in Fig. 6.20 - 6.24 and explained in the following.

In Fig. 6.20 and 6.21, the PESs are determined through the ionic intermediate $\text{INT1ION}_{(\text{AT})}$ and $\text{INT1ION}_{(\text{GC})}$, respectively, whose structures correspond to the LE-state optimized geometries (S_2 min). The atoms H_{61} in AT and H_1 in GC are first transferred from the purine to the pyrimidine moieties, therefore generating a charge separation between the monomers. A second proton, H'_3 in AT and H'_{41} in GC, is transferred from the pyrimidine moiety to the purine unit in order to connect the WC structure with the tautomeric form $\text{TAU}_{(\text{AT})}$ for the AT base pair and $\text{TAU1}_{(\text{GC})}$ for the GC dimer. It is worth noting that the state with CT nature (S_1) is more stable than the LE state (S_2) at the equilibrium structure of the latter around the INT1ION region. It evidences the relevance of the CT state in the radiationless deactivation pathways, as explained in the next mechanism.

In Fig. 6.22, the H'_3 proton is first transferred from T to A, obtaining the $\text{INT2ION}_{(\text{AT})}$ intermediate which is connected to $\text{TAU}_{(\text{AT})}$ through another proton transfer along the $O'_4 \cdots H_{61}N_6$ hydrogen bond. The S_1 PES is somewhat flat between the WC AT and $\text{TAU}_{(\text{AT})}$ species, indeed, the energy difference between the WC AT, $\text{INT2ION}_{(\text{AT})}$, and $\text{TAU}_{(\text{AT})}$ is found within an energy range of 0.28 eV. Therefore, the formation of the tautomeric form in the excited state PES via the SDPT mechanism is possible. Once in the S_1 excited state at the $\text{TAU}_{(\text{AT})}$ region, the system may decay via an internal conversion process through the CI of the 1,9H-6NH-adenine monomer.

In Fig. 6.23, the H'_{41} off-center proton is transferred initially from C to G, resulting in the $\text{INT3ION}_{(\text{GC})}$ intermediate. This structure is connected to

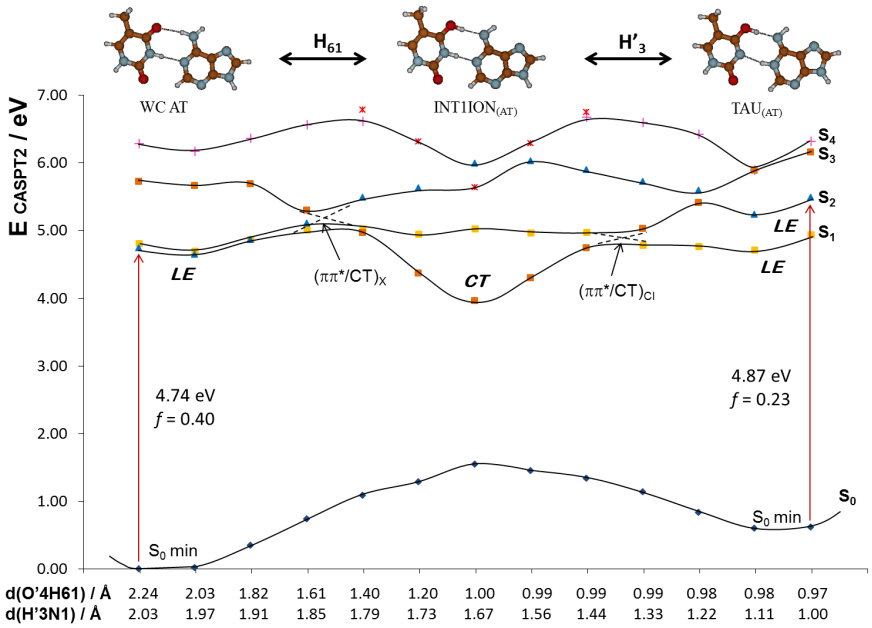


Figure 6.20: SDPT mechanism (H_{61}, H'_3). AT-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1, S_2, S_3, and S_4) are mapped at the CASPT2 level along the O'_4H_{61} and H'_3N_1 reaction coordinates between the WC AT and $TAU_{(AT)}$ structures via the $INT1ION_{(AT)}$ intermediate [equilibrium structure of the locally excited S_2 state (S_2 min)]. The computed stationary points for the ground (S_0 min) state are shown. The CASSCF geometry optimization of the S_3 state for the $INT1ION_{(AT)}$ species is the S_2 state at the CASPT2 level. The excited states show either charge transfer (CT), intra-adenine locally excited (AA), or intra-thymine locally excited (TT) character. To track the nature of the states, the dots on the graphics are drawn as follows: blue diamonds - ground state, blue triangles - ^{1}AA, yellow squares - ^{1}AA, orange squares - 1^1CT, red asterisks - 2^1CT, green circles - 3^1CT (when are shown), and pink crosses - ^{1}TT.

$TAU1_{(GC)}$ by means of a second proton transfer along the $N'_3\cdots H_1N_1$ hydrogen bond. As for AT, the S_1 PES is relatively flat between the WC GC and $TAU1_{(GC)}$ configurations. In this case, all the WC GC, $INT3ION_{(GC)}$,

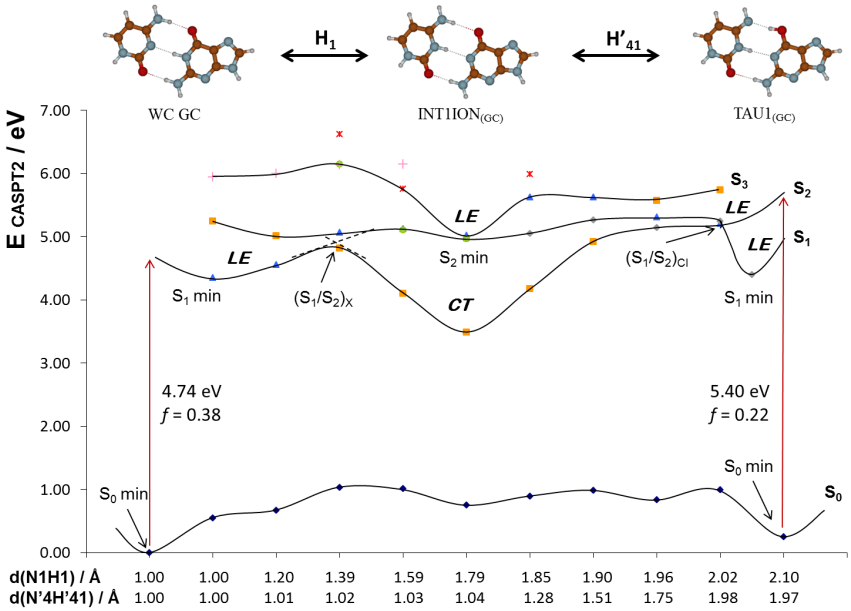


Figure 6.21: SDPT mechanism (H_1, H'_{41}). GC-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1, S_2, and S_3) are mapped at the CASPT2 level along the N_1H_1 and $N'_4H'_{41}$ reaction coordinates between the WC GC and TAU1$_{(GC)}$ structures via the INT1ION$_{(GC)}$ intermediate [equilibrium structure of the locally excited S_2 state (S_2 min)]. The computed stationary points for the ground (S_0 min) and excited (S_1 min and S_2 min) states are shown. The CASSCF geometry optimization of the S_4 state for the INT1ION$_{(GC)}$ species is the S_2 state at the CASPT2 level. The excited states show either charge transfer (CT), intra-guanine locally excited (GG), or intra-cytosine locally excited (CC) character. To track the nature of the states, the dots on the graphics are drawn as follows: blue diamonds - ground state, blue triangles - ^{1}GG, orange squares - 1^1CT, red asterisks - 2^1CT, green circles - 3^1CT, pink crosses - ^{1}CC, grey diamonds - ^{1}GG, yellow squares - ^{1}GG, and violet crosses - ^{1}CC.

and TAU1$_{(GC)}$ S_1 minima (S_1 min) are placed within an energy range of 0.41 eV. Hence, the SDPT mechanism allows the tautomer formation in the ex-

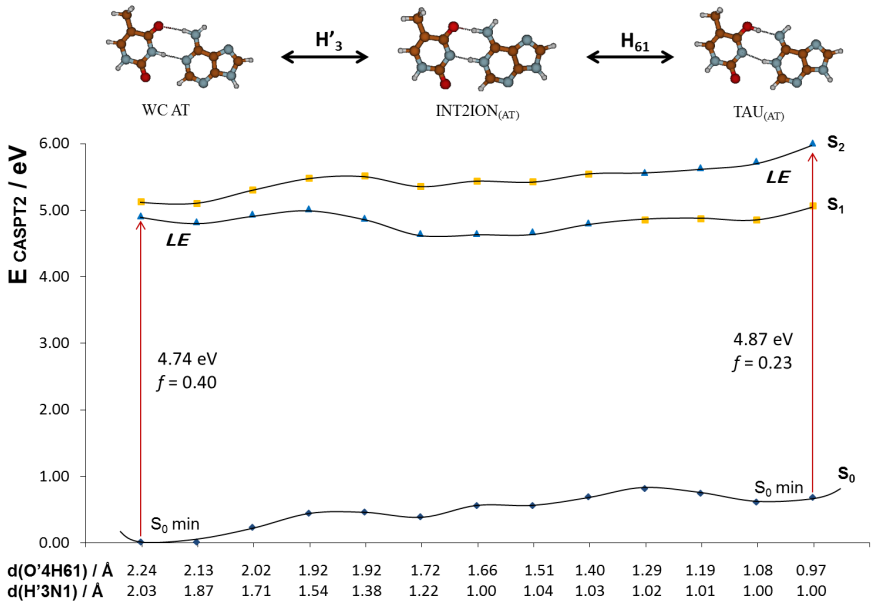


Figure 6.22: SDPT mechanism (H'_3, H_{61}). AT-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1 and S_2) are mapped at the CASPT2 level along the O'_4H_{61} and H'_3N_1 reaction coordinates between the WC AT and $TAU1_{(AT)}$ structures via the $INT2ION_{(AT)}$ intermediate [equilibrium structure of the locally excited S_1 state (S_1 min)]. The stationary PES points for the ground (S_0 min) state are shown. Color code follows that in Fig. 6.20.

cited state PES. In the region of the $TAU1_{(GC)}$ S_1 minimum, the system may decay nonradiatively through the CI of the 9H-6OH-guanine monomer. As previously described by Serrano-Andrés et al. [276], larger decay times are expected for this deactivation with respect to the natural G base due to the presence of barriers along the decay path in the former species not found in the canonical NAB.

In both AT and GC base pairs, the aforementioned tautomers ($TAU_{(AT)}$ and $TAU1_{(GC)}$) are relatively stable when comparing their ground-state equi-

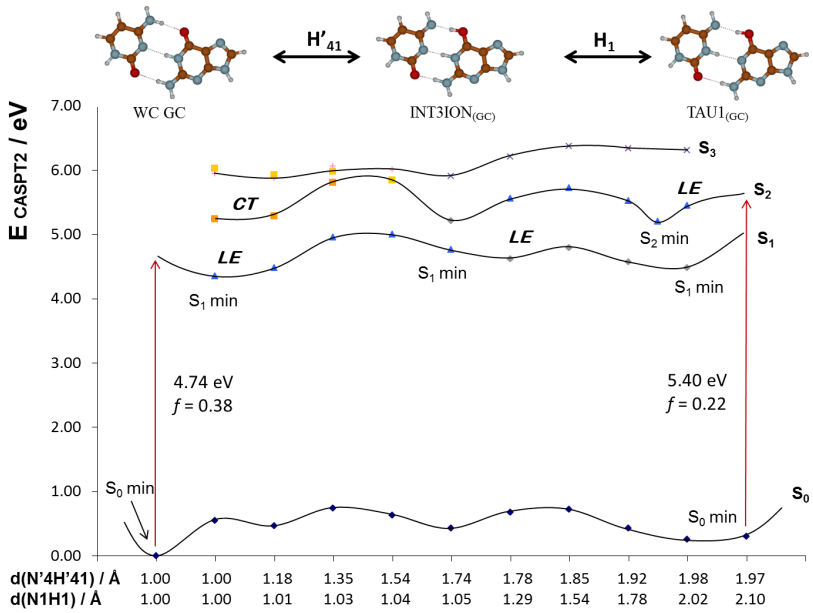


Figure 6.23: SDPT mechanism (H'_{41}, H_1). GC-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1, S_2, and S_3) are mapped at the CASPT2 level along the $N'_4H'_{41}$ and N_1H_1 reaction coordinates between the WC GC and $TAU1_{(GC)}$ structures via the $INT3ION_{(GC)}$ intermediate [equilibrium structure of the locally excited S_1 state (S_1 min)]. The stationary PES points for the ground (S_0 min) and excited (S_1 min) states are shown. Color code follows that in Fig. 6.21.

librium energies with their corresponding canonical base pairs (0.62 eV for $TAU_{(AT)}$ and 0.30 eV for $TAU1_{(GC)}$). It is worth mentioning that these tautomeric species do not cause DNA replication breakdowns. However, the tautomeric forms $TAU_{(AT)}$ and $TAU1_{(GC)}$ might lead to the GC and AT base-pair transition mutations, respectively [54].

In Fig. 6.24, the second tautomeric form ($TAU2_{(GC)}$) is connected with the $INT3ION_{(GC)}$ intermediate by means of a second proton transfer process from the amino group of G to the oxygen of C. The computed results for the S_1 PES

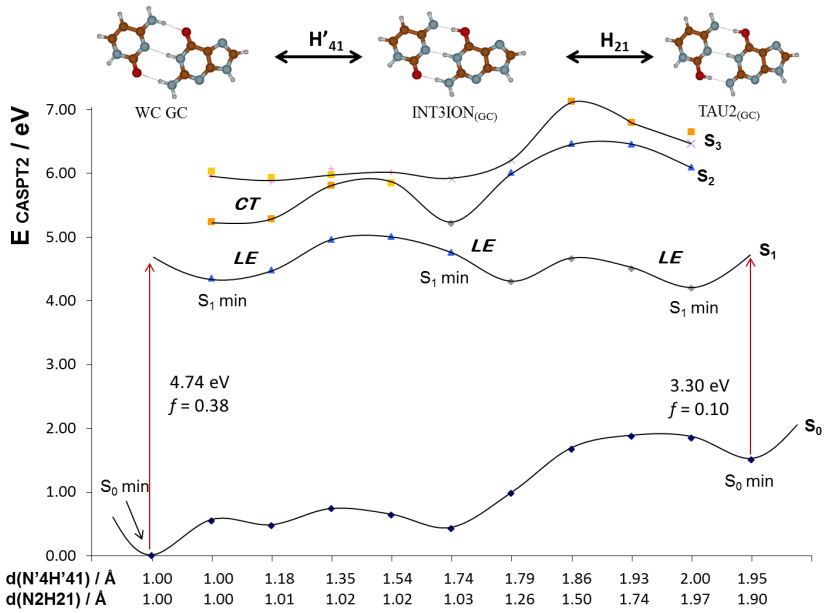


Figure 6.24: SDPT mechanism (H'_{41}, H_{21}). GC-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1, S_2, and S_3) are mapped at the CASPT2 level along the $N'_4H'_{41}$ and N_2H_{21} reaction coordinates between the WC GC and $TAU2_{(GC)}$ structures via the $INT3ION_{(GC)}$ intermediate [equilibrium structure of the locally excited S_1 state (S_1 min)]. The stationary PES points for the ground (S_0 min) and excited (S_1 min) states are shown. Color code follows that in Fig. 6.21.

also suggest the formation of tautomers in the excited state, the $TAU2_{(GC)}$ species. Similarly to $TAU1_{(GC)}$, a CI localized in the 1H-2NH-6OH-guanine monomer might also be expected to mediate a nonradiative energy decay of this modified NAB [276]. Conversely, this ground-state configuration of the GC is high in energy (1.50 eV) and low energy barriers are found on the ground-state PES, which makes the $TAU2_{(GC)}$ species unstable to thermal conversion toward the WC GC form (cf. Fig. 6.24). This property is highly important for the stability of the genetic code, since the formation of $TAU2_{(GC)}$ can stop DNA replication (the tautomeric monomers in $TAU2_{(GC)}$ cannot

match with canonical NABs, unlike those present in $TAU_{(AT)}$ and $TAU1_{(GC)}$) [54].

Stepwise double hydrogen-transfer (SDHT) mechanism

The key state within the basis of this mechanism has a CT nature and involves an excitation from the purine to the pyrimidine moieties. The charge separation between both monomers caused by the electronic excitation favors the proton transfer in the same direction. The initial hydrogen transfer takes place at the $O'_4 \cdots H_{61}N_6$ (Fig. 6.25) hydrogen bond in the AT base pair to form a neutral intermediate, $INT1NEU_{(AT)}$, and at the $N'_3 \cdots H_1N_1$ (Fig. 6.26) or $N'_4H'_{41} \cdots O_6$ (Fig. 6.27) hydrogen bonds in the GC base pair to give rise to neutral species, either $INT1NEU_{(GC)}$ or $INT2NEU_{(GC)}$, respectively.

The equilibrium structure of the CT state at the neutral intermediates appears in the region of near-degeneracy between the S_0 and S_1 PESs (see Fig. 6.25 - 6.27). Along the S_1 surface connecting the canonical and neutral-intermediate structures, the excited state changes its nature from AA or GG to CT passing the $(\pi\pi^*/CT)_X$ or $(S_1/S_2)_X$ avoided crossing characterized previously by Domcke and co-workers [81, 83]. From this structure, a barrierless pathway towards the crossing with the ground state, $(S_0/S_1)_{CI}$, is obtained, in agreement with previous results [81]. Both the avoided and conical intersection crossings imply an electron transfer process which is the driving force for the proton transfer process. In the first case, the electron is transferred from the purine to the pyrimidine moiety, and the charge separation between both NABs promotes a proton migration from the purine to the pyrimidine monomer as shown in Fig. 6.25 - 6.27. In the $(S_0/S_1)_{CI}$ crossing, the electron transfer occurs from the pyrimidine to the purine moiety allowing a second proton transfer in the same direction. The proton transferred can be either the same, hence opening a photostable channel in the AT and GC base pairs, as established previously, [81, 83] or other different proton migrating from the pyrimidine to the purine

moiety, hence forming the tautomeric forms. In the overall process two hydrogen atoms have been transferred.

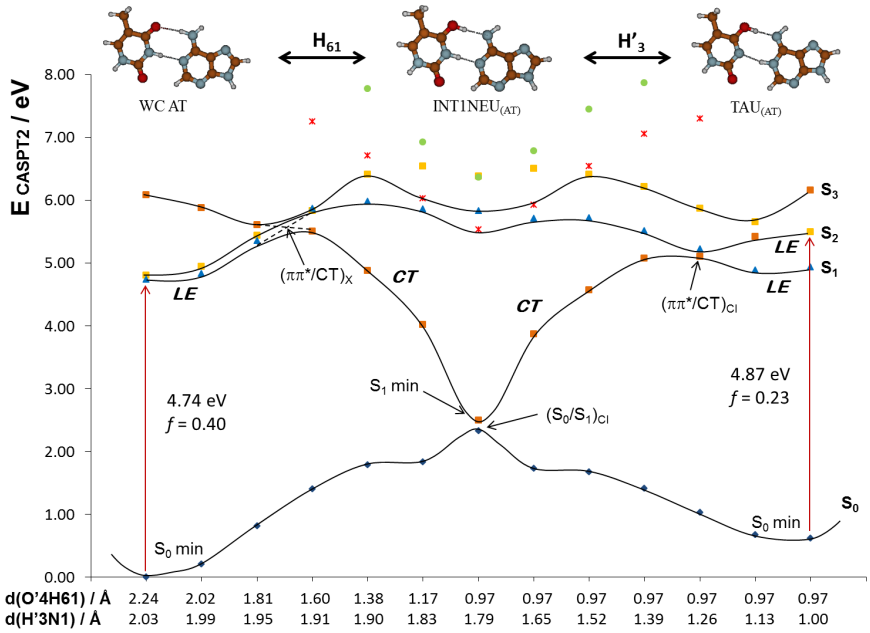


Figure 6.25: SDHT mechanism (H_{61}, H'_3). AT-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1, S_2, and S_3) are mapped at the CASPT2 level along the O'_4H_{61} and H'_3N_1 reaction coordinates between the WC AT and $TAU_{(AT)}$ structures via the $INT1NEU_{(AT)}$ intermediate [equilibrium structure of the charge transfer S_1 state (S_1 min)]. The computed stationary PES points for the ground (S_0 min) state are shown. Color code follows that in Fig. 6.20.

A novel contribution of this study is the mechanism established for the tautomerism of the DNA base pairs and how it is also photoreversible in the same manner as the canonical base pairs themselves. At the S_0 optimized geometry of the tautomers, the brightest low-lying $^1\pi\pi^*$ state has an intrapurine nature. Proton motions along the hydrogen bonds from the tautomeric purine to the tautomeric pyrimidine moieties can activate a barrierless nonadiabatic process from the brightest state to the S_1-CT state. Once in the

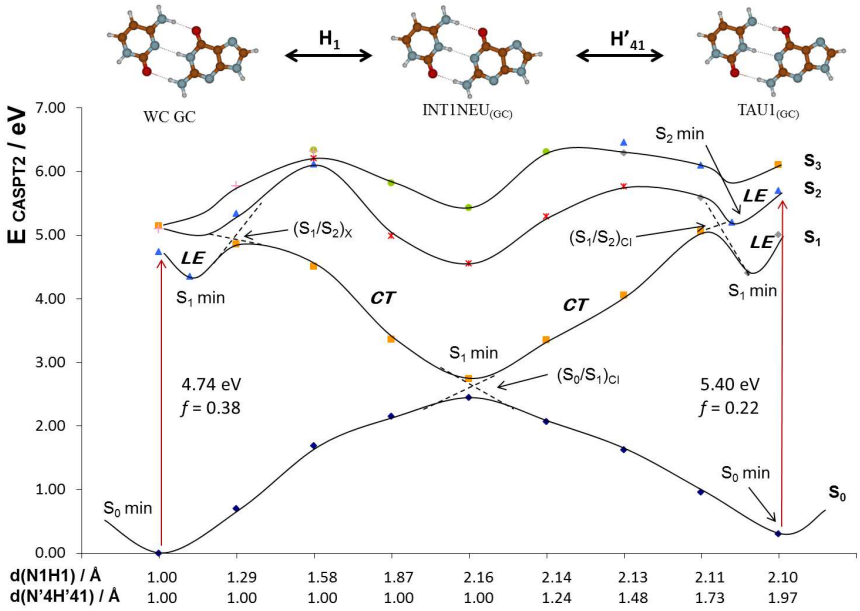


Figure 6.26: SDHT mechanism (H_1, H'_{41}). GC-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1, S_2, and S_3) are mapped at the CASPT2 level along the N_1H_1 and $N'_4H'_{41}$ reaction coordinates between the WC GC and $TAU1_{(GC)}$ structures via the $INT1NEU_{(GC)}$ intermediate [equilibrium structure of the charge transfer S_1 state (S_1 min)]. The computed stationary PES points for the ground (S_0 min) and excited (S_1 min and S_2 min) states are shown. Color code follows that in Fig. 6.21.

S_1 PES, the charge separation between both monomers forces H'_3 in $TAU_{(AT)}$ and H'_{41} in both $TAU1_{(GC)}$ and $TAU2_{(GC)}$ to be completely transferred from the tautomeric purines to the tautomeric pyrimidines, yielding the neutral intermediates. Thus, the $(S_0/S_1)_{CI}$ region which mediates the photostability of the WC base pairs is also reached in the case of the tautomers, funneling the energy to the ground-state equilibrium structure of either the tautomeric forms or the canonical base pairs. It is worth mentioning that the tautomer $TAU2_{(GC)}$ is photochemically less stable than $TAU1_{(GC)}$, in addition to the lower thermal stability explained previously.

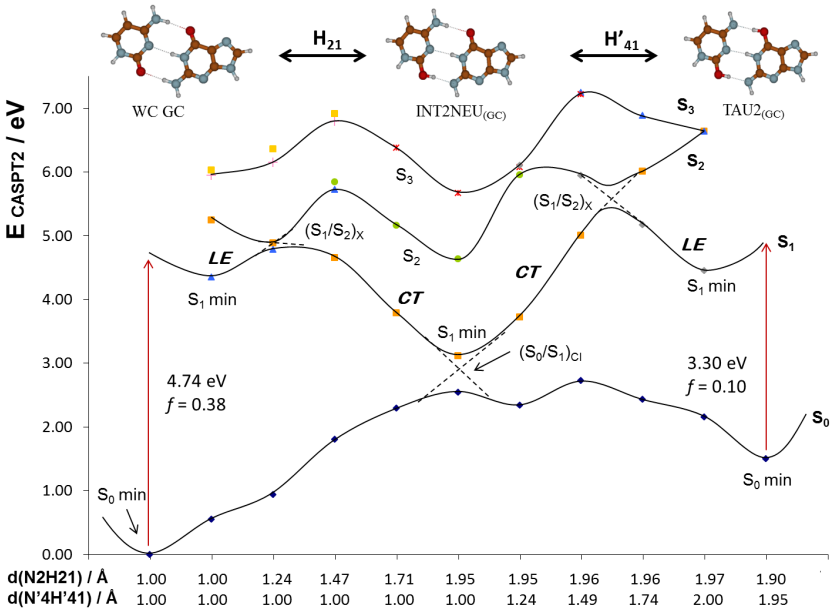


Figure 6.27: SDHT mechanism (H_{21}, H'_{41}). GC-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1, S_2, and S_3) are mapped at the CASPT2 level along the N_2H_{21} and $N'_4H'_{41}$ reaction coordinates between the WC GC and $TAU2_{(GC)}$ structures via the $INT2NEU_{(GC)}$ intermediate [equilibrium structure of the charge transfer S_1 state (S_1 min)]. The computed stationary points for the ground (S_0 min) and excited (S_1 min) states are shown. Color code follows that in Fig. 6.21.

In general, the excited canonical and tautomeric species in both AT and GC base pairs are interconnected in the SDHT mechanism via a nonadiabatic process to their ground states. This points out to a non-unique ultrafast nonradiative decay channel for the WC configurations, in contrast to previous proposals [81, 82, 84]. Therefore, ultrafast conversions between the WC and tautomeric forms can be expected during the UV irradiation of the DNA base pairs.

Concerted double proton-transfer (CDPT) mechanism

SDPT and SDHT mechanisms involve two steps in which one proton/hydrogen per step is exchanged between the NABs in the base pair via an intermediate. The synchronous proton/hydrogen transfer along the hydrogen bonds also connects the WC configurations with the tautomeric forms of the DNA base pairs (see Fig. 6.28 - 6.30). However, high energy barriers are found on the excited state PESs, which makes the concerted mechanism clearly unfavorable for the UV-activated double proton transfer.

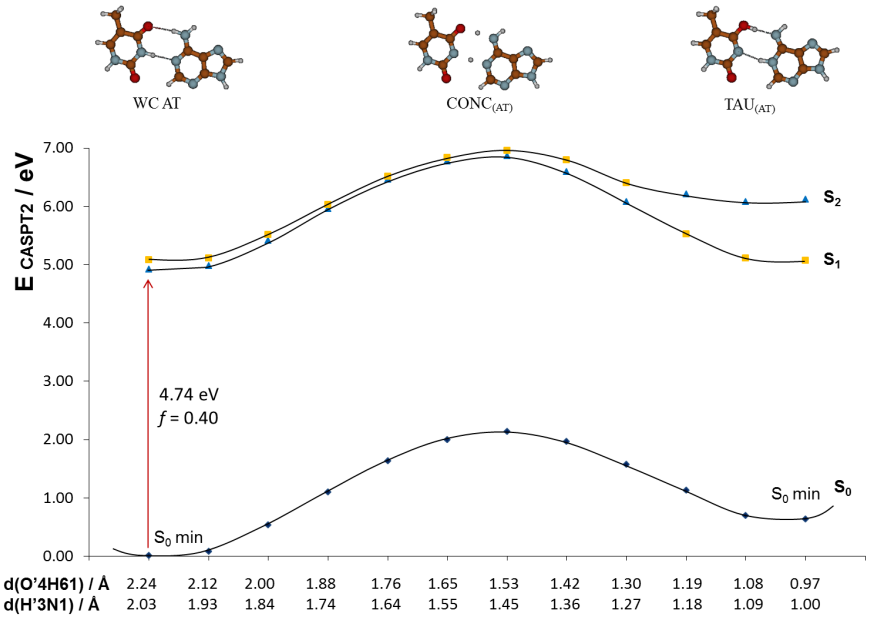


Figure 6.28: CDPT mechanism (H_{61}+H'_3). AT-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1 and S_2) are mapped at the CASPT2 level along the O'_4H_{61} and H'_3N_1 reaction coordinates between the WC AT and $TAU_{(AT)}$ structures via the concerted pathway. The computed stationary points for the ground (S_0 min) are shown. Color code follows that in Fig. 6.20.

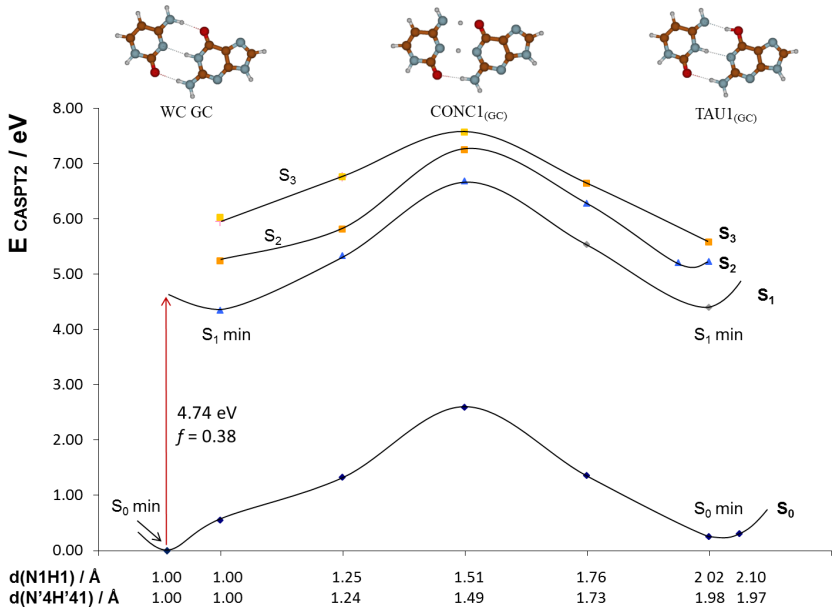


Figure 6.29: CDPT mechanism (H_1+H'_{41}). GC-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1, S_2, and S_3) are mapped at the CASPT2 level along the N_1H_1 and $N'_4H'_{41}$ reaction coordinates between the WC GC and TAU1$_{(GC)}$ structures via the concerted pathway. The computed stationary points for the ground (S_0 min) and excited (S_1 min) states are shown. Color code follows that in Fig. 6.21.

In addition, ground-state PESs show larger activation barriers, which reduce the significance of the thermal conversion process, in favor of the SDPT mechanism (cf. Fig. 6.28 - 6.30 vs. Fig. 6.20 - 6.24).

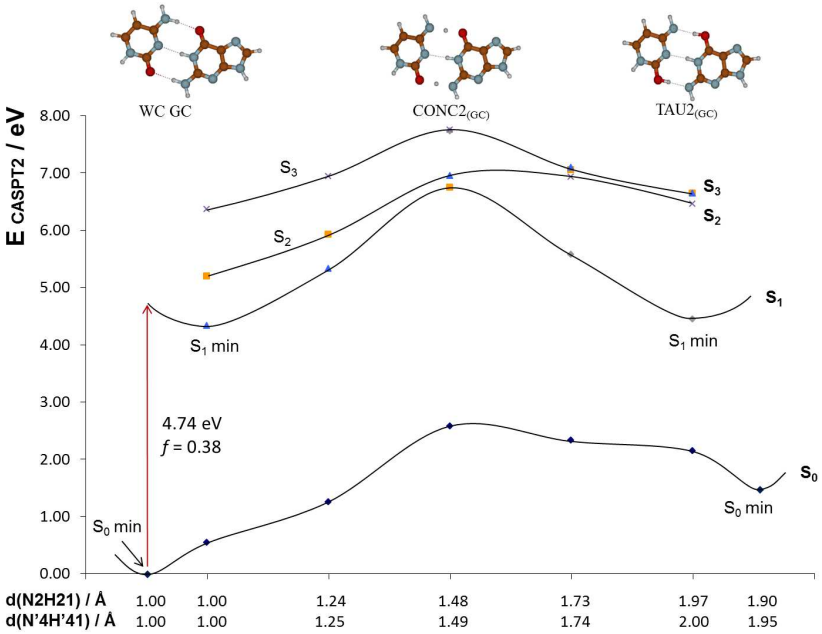


Figure 6.30: CDPT mechanism (H_{21}+H'_{41}). GC-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1, S_2, and S_3) mapped at the CASPT2 level along the N_2H_{21} and $N'_4H'_{41}$ reaction coordinates between the WC GC and $TAU2_{(GC)}$ structures via the concerted pathway. The stationary PES points for the ground (S_0 min) and excited (S_1 min) states are shown. Color code follows that in Fig. 6.21.

6.4.2. GC base pair in the biological environment

SDHT has been established in the preceding section as an efficient mechanism for nonradiative energy decay toward either the canonical GC or the tautomeric form. The process involves states of CT nature and ionic species which are sensitive to solvent interactions. In fact, a clear energy separation between the ground and lowest-lying LE and CT states along the N_1H_1 reaction coordinate has been computed by Biemann et al. [92] in a recent TD-DFT study on the GC base pair in solution with $CHCl_3$. This is fully consistent with

the fact that the ground state is ionic at the geometry of the intermediates, whereas S_1 is neutral. Based on their findings, Biemann et al. concluded that the hydrogen transfer is not efficient in $CHCl_3$ and cannot be considered as a relevant deactivation route in DNA [92]. However, the conditions in which NABs are present in nucleic acids (inner part of the double strand, surrounded by the sugar-phosphate backbone) are different from solution and comparisons are not straightforward.

To clarify these aspects, the differences between the SDPT and SDHT mechanisms are analyzed in the gas phase and in a DNA-embedded GC base pair. The main structures of the two mechanisms obtained *in vacuo* have been considered in this study. Hence, the equilibrium structures of the WC, INT1NEU, INT2NEU, INT3ION, TAU1, and TAU2 species (where the label GC has been dropped) in a $d(GC)_9 \cdot d(GC)_9$ strand in aqueous solution have been determined at the CASSCF/MM level of theory. Figure 6.31 illustrates the general scenario for the SDPT and SDHT channels in the DNA, compiling the CASPT2 energies for the ground (S_0) and lowest-lying excited (S_1) states, together with the corresponding values in the gas phase. The lowest-lying excited state at the DNA-embedded WC GC base pair corresponds to a locally excitation (LE), as in the isolated base pair. DNA vertical energies are in general close to the gas phase results, which points out to similar properties of the SDPT and SDHT mechanisms in both gas and DNA phases. Crucial structures in the SDHT mechanism are the intermediates INT1NEU or INT2NEU, which are located in the region of near-degeneracy between the S_0 and S_1 PESs. The energy gap of 0.3 eV, found *in vacuo* for the INT1NEU intermediate, is preserved in the DNA, whereas the INT2NEU equilibrium structure is even closer to the CI in the biological environment. The INT3ION species, which has an ionic structure, is slightly stabilized in DNA. However, such stabilization is not sufficient to facilitate an internal conversion process via the SDPT mechanism. In summary, the GC base pair in both the gas phase and in DNA manifests similar attributes for the proton/hydrogen transfer processes.

In the light of the present QM/MM results, it is found that the excited-state proton/hydrogen process in the GC base pair seems to be a plausible

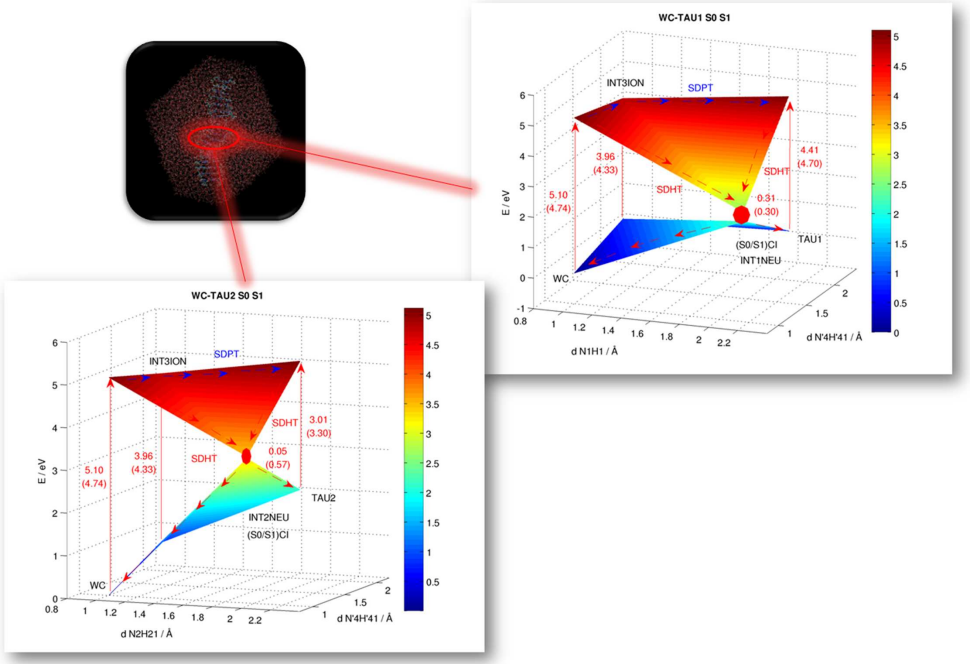


Figure 6.31: Stepwise double proton-transfer (SDPT) and stepwise double hydrogen-transfer (SDHT) mechanisms in the DNA-embedded GC base pair [d(GC)$_9$·d(GC)$_9$] between the canonical WC form and the tautomeric TAU1 (top) and TAU2 (bottom) species. DNA CASPT2 vertical energies (in eV) for the lowest-lying excited state (S_1) are shown, together with the related gas-phase values (within parentheses). S_1 corresponds to a locally excited state (LE) in the WC, INT3ION, and TAU1 structures and to the charge transfer state (CT) in the INT1NEU and INT2NEU systems. In TAU2, the nature of S_1 differs from the results *in vacuo*.

photochemical channel for the actual photostability of DNA and the production of tautomers. The SDHT energy-decay pathway will compete with the monomeric nonradiative photochemical channel, also present in the isolated base pair, and with intrastrand phenomena, such as the formation of exciplexes between π-stacked NABs, as suggested by experiments in alternating and non-alternating GC oligonucleotides [86, 88] and previous computations on related

systems [280, 281]. The efficiency of the hydrogen/proton energy-decay pathways in DNA will depend on the degree of hydrogen-bonding between the strands.

6.4.3. Excited-state dynamics of the GC base pair

In order to evaluate some qualitative attributes related to the different energy-decay channels, *post-vacuo* and DNA computations including dynamical and statistical aspects are next considered. Previous works on the dynamics of the system have not been able to achieve the correct electronic structure of the states at the FC region and the population of the lowest-lying excited state with CT character is enhanced as an artifact of the low level of theory used [89]. Therefore, those earlier computations provide a qualitative incomplete description of the overall photochemistry. In the present study, four dynamics simulations have been performed employing the methodology used to describe the proton/hydrogen transfer mechanisms *in vacuo* and in the DNA to get a qualitative insight into the dynamical properties of the proton/hydrogen transfer phenomenon in the GC base pair, leaving for the future a full study able to compute quantitative quantum yields for the formation of the products.

Table 6.15 compiles the values of the $N'_4H'_{41}$, N_1H_1, and N_2H_{21} bond lengths and projected scalar velocities of the H'_{41}, H_1, and H_{21} atoms over the lines defined by the three hydrogen bonds for four initial conditions which have been selected, among a Wigner distribution of 300 points, to perform the dynamics simulations. The selected conditions have appropriate NH distances and H momenta for the analysis of some dynamical attributes of the different relaxation mechanisms of the GC base pair (radiationless decay via the purine moiety, hydrogen/proton transfer via the SDPT or SDHT mechanisms, and formation of different tautomers). Figure 6.32 displays the results obtained for the magnitude of the dipole moments ($|\mu|$) of the ground (S_0) and low-lying excited (S_1) states, the $N'_4H'_{41}$, N_1H_1, and N_2H_{21} bond lengths, and the S_0, S_1, and S_1-S_0 energies.

Table 6.15: NH bond lengths and scalar projections of the velocities of the H'_{41}, H_1, and H_{21} atoms over the lines defined by the $N'_4H'_{41}\cdots O_6$, $N'_3\cdots H_1N_1$, and $O'_2\cdots H_{21}N_2$ hydrogen bonds, respectively, in the GC base pair. The mechanisms and photochemical decay paths observed in the simulations are indicated.

Simulation	bond length / Å			projected velocity / au			mechanism	photochemistry
	$N'_4H'_{41}$	N_1H_1	N_2H_{21}	H'_{41}	H_1	H_{21}	activated	
Sim1	1.26	0.92	0.99	0.0032	0.0005	-0.0006		nonradiative decay via CI of G
Sim2	1.20	1.02	0.99	0.0028	0.0030	-0.0004	SDHT	photostability GC/TAU1 formation
Sim3	0.97	0.87	1.14	0.0008	0.0006	0.0041	SDHT	photostability GC/TAU1 formation
Sim4	1.05	1.07	1.20	0.0017	0.0042	-0.0035	SDHT	photostability GC/TAU1 formation

These properties allow us to identify the relaxation mechanism which takes place in each one of the simulations. In particular, $|\mu|$ can be used to analyze the electronic structure of the system along the simulations. At the FC region, large values ($|\mu| > 15$ D) indicate a CT nature of the states, whereas low values ($|\mu| = 0$ - 15 D) correspond to LE excitations. This behavior is inverted at the intermediates region in which a proton is exchanged between the NABs. Meanwhile, the NH bond distances and the energy difference between the S_0 and S_1 states give information on the region of the PESs where the system is located.

The dynamics *simulation* Sim1 features small momenta along the $N'_3 \cdots H_1N_1$ and $O'_2 \cdots H_{21}N_2$ hydrogen bonds and a large value at the $N'_4H'_{41} \cdots O_6$ site, which favors initially the formation of the excited state tautomeric forms via the SDPT mechanism (see Fig. 6.23 and 6.24). However, along the simulation, the system does not show any proton/hydrogen exchange between both NABs in the base pair. The CT nature of S_1 is reached at around 40 fs (large dipole moment in Fig. 6.32) due to a slight increase in the N_1H_1 bond distance. Subsequently, the system comes back to the region of LE character where is relaxed towards the equilibrium structure of this LE state. From this point, the ultrafast decay path through the CI localized in the G monomer might take place [276]. This decay pathway has not been found in the trajectories performed by Groenhof et al. [89], due to the fact that all the simulations were started in the CT state, as a result of the lower level of theory employed. *Simulation* Sim2 includes additionally an initial velocity component in the direction of the $N'_3 \cdots H_1N_1$ hydrogen bond, favoring also the radiationless energy decay toward the TAU1 species or the restored WC base pair, via the SDHT mechanism (Fig. 6.26). In this case, the GC base pair has a different fate. The system, initially on the S_1-LE state, evolves in a few femtoseconds into the S_1-CT state. At 25 fs, an inversion of the dipole moments of the ground and excited states takes place. Simultaneously, the H atom is transferred from G to C. Hence, the INT1NEU structure is reached, although the relaxation toward the ground state surface would not be still possible since the energy difference between the states is large. The $(S_0/S_1)_{CI}$ crossing is reached around 40 fs.

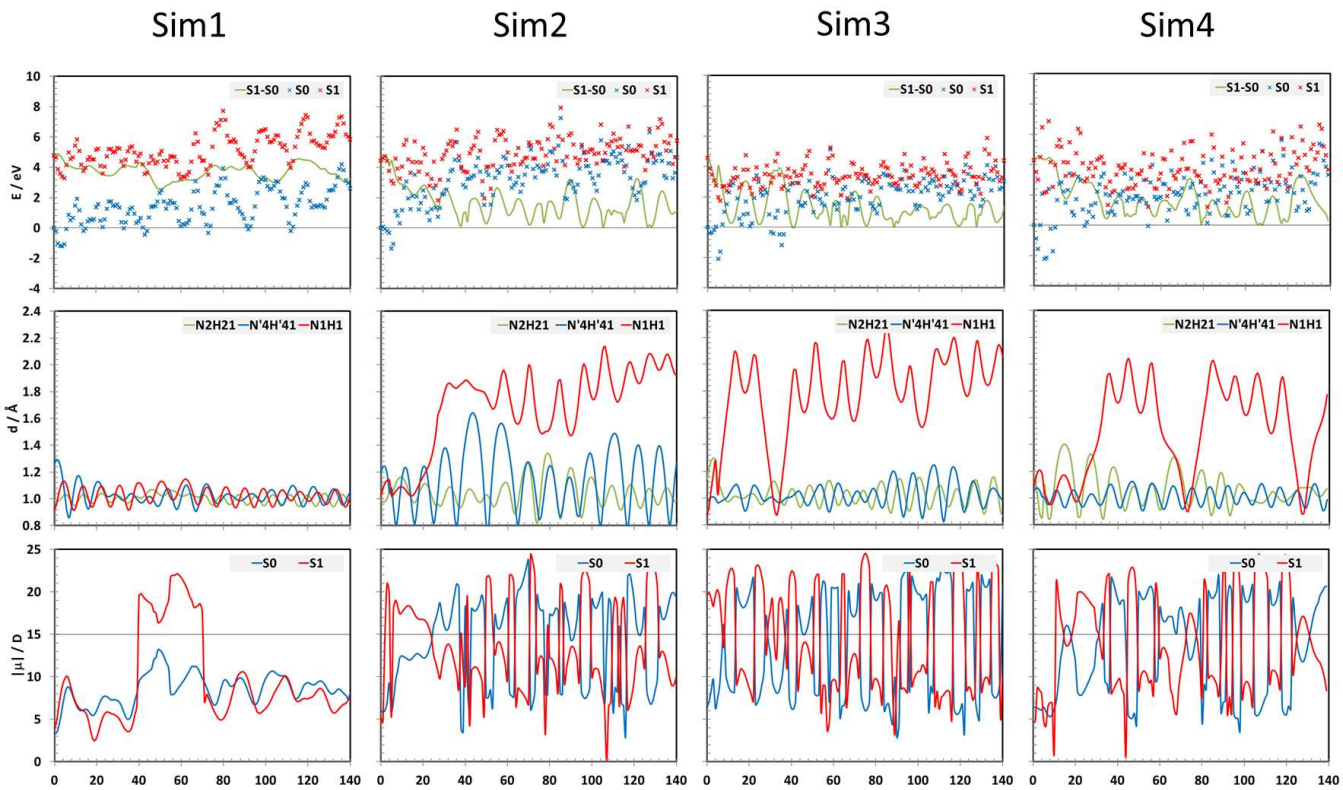


Figure 6.32: CASPT2 energies (E) for the ground state (S_0), the lowest-lying excited state (S_1), and energy difference between both states (S_1-S_0); NH distances (d); and the magnitude of the dipole moments ($|\mu|$) of the S_0 and S_1 states vs. time (t) are plotted for the semiclassical dynamics simulations Sim1, Sim2, Sim3, and Sim4.

It is worth noting that after this point the system is trapped in the INT1NEU structure and there is a high probability for internal conversion. Groenhof et al. obtained a recovery of the WC configuration in 75% of 20 simulations [89]. In five out of 20 runs, the TAU1 configuration was found. *Simulation* Sim3 points to the ground-state formation of the TAU2 species, instead of TAU1, via the SDHT mechanism (Fig. 6.27). The calculation starts already with a CT character of the excited state (see dipole moment for S_1 in Fig. 6.32). Furthermore, although that a large momentum is located initially in the $O_2' \cdots H_{21}N_2$ direction pointing toward the C moiety, the transfer of H_1 from G to C is again taking place. The GC base pair reaches the INT1NEU configuration in a few femtoseconds, after the first 40 fs and a third hydrogen transfer at the center hydrogen bond, the system is trapped in the well of the INT1NEU. Large changes in the CASSCF and CASPT2 energies for the S_0 and S_1 states are found initially, and later the S_1-S_0 energy difference becomes small. The fourth calculation performed on the dynamics of the GC base pair after UV irradiation, *simulation* Sim4, is characterized by significant initial momenta of the H atoms in all the three hydrogen bonds. During 10 fs, both ground and excited states have low dipole moment, meaning a trapping of the system in the LE region. Later, in the GC base pair the evolution toward CT nature takes place, a large increase of the N_2H_{21} bond length appears, and the energy difference between the S_0 and S_1 states drops significantly. The SDHT mechanism via the INT2NEU intermediate seems to be activated. In fact, the ground and excited state dipole moments show an inversion. Once the N_2H_{21} distance reaches a maximum value of 1.4 Å, the dimer distributes the kinetic energy among the other vibrational modes and finally the SDHT mechanism via the INT1NEU structure is turned on instead. At 35 fs, all the control properties, i.e. magnitudes of the dipole moments, N_1H_1 bond length and vertical relative energy between S_0 and S_1 states, point simultaneously to the formation of the INT1NEU intermediate. Despite the simulation shows an oscillatory pattern in the hydrogen transfer, there is a window at 30-60 fs, where the system might exhibit a nonadiabatic process toward the ground state PES.

In general, the hydrogen transfer at the $N_3' \cdots H_1N_1$ hydrogen bond is the most favorable mechanism in the four simulations carried out. The formation of the INT2NEU intermediate, which involve a hydrogen transfer at the $O_2' \cdots H_{21}N_2$ position, has not been obtained. INT1NEU requires less than 50 fs to be generated, which is a much faster process than the energy deactivation along the flat PES of the G NAB up to the methanamine-like CI, previously estimated around 100 fs [276]. This difference in the lifetime of both processes is probably a consequence of the fact that the former phenomenon involves the motion of mainly one H atom and that the PES is steep, while the energy relaxation toward the CI of the monomer implies out-of-plane distortions of the six-member ring along a flat PES. Tunneling effects, not considered in the employed semiclassical approach, can be expected to decrease the lifetimes. Nevertheless, the present results must be used with caution when comparing with the experiments. Although lifetimes shorter than the energy-decays in the G and C mononucleotides were measured by means of fluorescence up-conversion techniques [86], de La Harpe et al. [88] obtained energy-decays an order of magnitude larger with respect to the building blocks. This suggests the formation of relatively stable excited-state species by means of π-stacking interactions, which are not modeled here. In the present work, at the time of UV irradiation of the GC base pair, those molecules with a certain momentum in the $N_3' \cdots H_1N_1$ direction (for instance, Sim2, Sim3, and Sim4) might undergo a double hydrogen transfer process, resulting in the production of the TAU1 configuration or restoring the WC structure. On the other hand, other samples with low momenta in the hydrogen atoms from the hydrogen bonds (for instance, Sim1) will relax the energy through the CI crossing characterized in the G monomer, as pointed out in previous studies on the isolated molecule [276], or form relatively stable exciplexes via the π-stacking network [86].

Chapter 7

Conclusions

In the present thesis, several works of photochemical and photophysical interest have been carried out yielding attractive contributions to the scientific community. The main concluding remarks are presented in four different parts which correspond to the four important studies developed during this predoctoral investigation period: The RASPT2 method to compute electronic excited states, the parameterization of the extended Hubbard model for conjugated and aromatic hydrocarbons derived from stretched ethene, the distinct photophysics of the $B_{18}H_{22}$ isomers, and finally, the proton/hydrogen transfer mechanisms in the Adenine-Thymine and Guanine-Cytosine base pairs to explain photostability and tautomerism in DNA.

The RASPT2 method for electronic excited states

The electronic excited states of a set of organic (free base porphin, ethene, benzene) and inorganic (the nickel atom and the copper tetrachloride dianion) systems have been computed at the RASSCF/RASPT2/MS-RASPT2 levels

of calculation using different active space partitions and strategies. The results have been compared to those obtained with well-established procedures like CASSCF/CASPT2/MS-CASPT2 or CCSD, and to experimental data, in order to determine the accuracy of several procedures to choose the RAS space.

Free base porphin has been first investigated with several partition procedures. RASPT2 has proved to be an excellent strategy to avoid arbitrary divisions of the π space in a system in which the full-π active space (26 electrons in 24 MOs) is out of reach for the CASPT2 method. It has been shown that in the RASPT2 method the proper definition of the RAS2 space (in which a full CI is performed to define the configurational reference space) is crucial to assess the accuracy of the calculations. When computing a RASPT2 energy difference, the highest accuracy is obtained when the MOs changing the most their occupation number from one state (typically the ground state) to the other (an excited state) are all included in the RAS2 space, leaving the other less significant MOs in the RAS1/RAS3 spaces. If this requirement is fulfilled, a single-double (SD) level of excitations in these two latter active spaces partitions is sufficient to get a high accuracy. In free base porphin, as it is typical in many other organic systems, Gouterman's four MOs (HOMO, HOMO-1, LUMO, LUMO+1) form the basic set required to describe the four low-lying $\pi\pi^*$ excited states, and therefore it will be sufficient to include them in RAS2 while leaving the remaining $\pi\pi^*$ MOs in RAS1/RAS3 and reaching a SD level of excitation to get accurate results. Higher states will, however, require extension of the RAS2 space to include additional MOs. If the required MOs are excluded from RAS2, the results would become unbalanced in the CI treatment, and the second-order perturbation correction may not be able to compensate the results. Particularly for this case, a SD level of excitation is clearly insufficient. It is possible to improve the results, even if some important MOs are excluded from RAS2, when using different levels of excitation for the two considered states, like SD for the ground and SDT for the excited state, or increasing the overall excitation as much as possible, SDT, or even better, SDTQ, although these latter strategies might be impossible to apply because of the very large configurational spaces. All these results open the possibility

to use RASPT2 for many organic systems with extended π spaces, without further loss of accuracy due to restrictions in the size of the active space.

Calculations on the valence and Rydberg singlet and triplet excited states of ethene and benzene have been performed in the RASPT2 method. A new strategy for the RAS partition has been used in which the Rydberg MOs - typically nine ($n = 3$) for common organic systems - are placed in RAS3, leaving in the RAS2 space the valence $\pi\pi^*$ MOs and electrons, and allowing, apart from the full CI expansion within RAS2, just single excitations toward RAS3. The Rydberg orbitals can be moved out of the RAS2 space and, therefore, the computational effort is substantially decreased. The calculations are simpler when comparing to CASPT2 analogues because it is now feasible the use of a unique active space for the different symmetries. This approach, however, does not solve directly the valence-Rydberg mixing problems already found in CASSCF/CASPT2, leading to too high excitation energies and heavily mixed wave functions with too large orbital extensions for some valence states. As previously shown, when only the $\pi\pi^*$ MOs are included in the RAS2 space, the multi-state (MS) procedure, MS-RASPT2, is required to solve the mixing and provide orthogonal states with clear valence or Rydberg mixings. In the ethene case, the inclusion of the $\sigma\sigma^*$ MOs in the RAS1 and RAS3 spaces (not possible in general for CAS calculations) determines that the RASPT2(SD) level of calculation is not sufficient to correctly describe the $\sigma\sigma^*$ excitations. Increasing the excitation level to SDT solved the problem in the ethene case, although this might not be a general rule. When including both π and σ correlation within the CI reference space, the valence-Rydberg mixing was solved at the RASPT2(SDT) level, without using the MS-RASPT2(SDT) procedure. This shows the importance of electronic correlation in defining the wave function when dealing with the valence-Rydberg mixing problem.

Regarding the calculation of the first-row transition metal systems, the interest is related to the effect on the excitation energies of moving the $4d$ correlating shell from RAS2 to RAS3. The electronic spectra of the nickel atom and the copper tetrachloride dianion have been analyzed. The RASPT2 calculations quite well reproduce the corresponding CASPT2 results (to within 0.1-0.2 eV),

although a few exceptional cases were also observed, e.g. the $^1S(3d^{10})$ state of the nickel atom.

Parameterization of the extended Hubbard model for conjugated and aromatic hydrocarbons derived from stretched ethene

A simple extension of the Hückel model, adding the one-center electron-electron repulsion as described by the Hubbard Hamiltonian, leads to an attractive model of chemical bonding. It has been shown that the extended form of the Hubbard model, including also nearest-neighbor electron-electron interactions, can be parameterized in a readily transferable form from accurate potential curves for a two-site fragment. Accurate potential energy curves for ethene have been calculated via the MS-RASPT2 method. A multireference method such as MS-RASPT2 is essential to account for valence-Rydberg mixing in the $\pi\pi^*$ excited states of ethene as well as many other small molecules. The resulting potential curves have been used to extract parameters for the Hubbard model as a function of bond length, which should be applicable to other conjugated and aromatic hydrocarbons as well as pure carbon species such as graphene.

In calibration calculations on butadiene, hexatriene, and benzene the Extended Hubbard model has been found to give a reasonable description of carbon-carbon bond lengths, intermediate between equilibrium bond lengths calculated at the MP2 level and experimental bond lengths. MS-RASPT2 calculations have been carried out for these molecules and the Hubbard model was found to provide in general a good description of low-lying electronic excitation energies. That the model does not accurately reproduce the ^1B_u excited state in butadiene and hexatriene illustrates a limitation of all nearest-neighbor models. The simple Hubbard model is reliable only for covalent states, that is in neutral states dominated by doubly excited configurations. The extended form of the model used here should correctly describe ionic states provided that the charges are on adjacent sites, which is not the case in the ^1B_u states

of polyenes. To position them correctly longer-range Coulomb interactions are needed.

Within the limitation of the description of ionic states in polyenes, the Extended Hubbard model successfully includes the most important interactions required to understand chemical bonding. This model provides a more balanced description of π versus σ bonding energy, and it is capable of simultaneously describing geometric structure and low-level electronic excitations.

The full power of the model comes from the inclusion of electron correlation. For small molecules, such as those considered in this work, exact (full configuration interaction) solutions are feasible, and for larger molecules any of the standard quantum chemical methods for including electron correlation can be applied but at much reduced cost compared to an *ab initio* calculation.

Distinct photophysics of the $B_{18}H_{22}$ isomers

The combined experimental and computational work on the photophysics of the $B_{18}H_{22}$ isomers rationalizes the known observation that the *anti* isomer displays an intense fluorescence, whereas its *syn* isomer shows no such activity. The measured lowest-lying singlet excited state S_1 lifetime in *anti*-$B_{18}H_{22}$ is two orders of magnitude longer than in the corresponding *syn* isomer, and the latter decays nonradiatively on a picosecond timescale. Quantum-chemical computations based on the CASPT2//CASSCF protocol point out that the difference in the geometries of the two isomers at the conical intersection crossings along their S_0 and S_1 PEHs is the key determining factor in their different decay processes. Thus, the *syn* isomer shows one very large B-B elongation of about 1.2 Å at the CI, whereas the *anti* isomer has B-B elongations smaller than 1.0 Å at its corresponding $(S_0/S_1)_{CI}$. The absorbed energy in S_1 for *anti*-$B_{18}H_{22}$ is therefore redistributed vibrationally into several regions of the molecule rather than almost completely into a single vibrational mode as in

the case for the radiative *syn* isomer. The consequent prolonged S_1 lifetime for the *anti* isomer allows for relaxation via fluorescence.

The interest of the present work is related to the synthesis of derivatives of the $B_{18}H_{22}$ isomers with cluster substituents that do not alter the overall cluster geometries of the boron hydride compounds in order to obtain efficient spin-orbit coupling between singlet and triplet excited states that eventually give rise to singlet-oxygen production.

Proton/Hydrogen transfer mechanisms in the Adenine-Thymine and Guanine-Cytosine base pairs: photostability and tautomerism

The ultrafast deactivation mechanisms in the isolated nucleobases and the formation of lesions in π-stacking pyrimidine dimers explain, together with the results on the DNA base pairs collected in this thesis, the photochemistry and photoreactivity of DNA.

Three different mechanisms have been introduced in the present thesis as possible candidates for energy decay in the DNA base pairs: 1) stepwise double proton transfer, SDPT, 2) stepwise double hydrogen transfer, SDHT, and 3) concerted double proton transfer, CDPT. Based on the present findings, SDHT is potentially the most favorable mechanism to explain the photostable properties of the canonical DNA base pairs and also the formation of tautomers. In the SDHT mechanism, proton motions from the purine to the pyrimidine moieties favor the evolution of the lowest-lying excited state from LE to CT character, which subsequently drives a hydrogen transfer to form the corresponding neutral intermediates (Fig. 6.25 - 6.27). The latter species are located in a region of near-degeneracy with the ground-state surface, which can funnel the reaction path toward the restored WC DNA base pair or a tautomeric new form via a second hydrogen transfer at the same or a different inter-base-pair hydrogen bond, respectively. Analogous SDHT processes are applied to the tautomeric forms, resulting in the canonical base pairs or

restoring the tautomers. It is worth noticing that the C-based imino-enol tautomer in Fig. 6.27 is a high-energy configuration among the GC base pairs studied, with the lowest energy barrier to restore the WC form. According to the energy profiles obtained (Fig. 6.20 - 6.30), thermal conversions are more favorable via the SDPT mechanism, as compared to the SDHT and CDPT mechanisms.

In solution with polar solvents, the SDHT mechanism in GC is presumably altered inactivating the internal conversion channel which brings the WC GC base pair back to its ground-state equilibrium structure or to new tautomeric configurations [92]. Nevertheless, the DNA-embedded GC dimer studied by the QM/MM computational approach manifests a similar behavior relative to the isolated base pair for the photochemical energy-decay pathways involving double proton/hydrogen transfer, which supports experimental observations [86, 87] and previous theoretical predictions [89] of the photoinduced proton/hydrogen transfer process in DNA duplexes. Hydrogen-transfer phenomena in GC seem to be more feasible along the center hydrogen bond. Lifetimes for the hydrogen transfer (around 50 fs) are found to be much shorter than the nonradiative decays localized on the monomers. A direct comparison with the excited-state lifetimes measured in the experiments is not possible, since intra-strand photochemical processes, such as the formation of exciplexes, might affect the lifetimes.

In conclusion, new photochemical routes are accessible in the UV-irradiated DNA base pairs, which will compete with the energy-decay pathways of the isolated NABs and π-stacked (or intra-strand) dimers. The former processes will be more efficient in highly hydrogen-bonding conformations of the DNA double strand. Vibrations at the hydrogen bonds can activate hydrogen/proton transfers in the DNA base pairs, as suggested by the simulated dynamics of the GC dimer (Tabla 6.15). The overall process is expected to be a double hydrogen transfer, which provides the system with new photostable routes, in addition to those present in the isolated NABs. The formation of tautomers (DNA photoreactivity) is also possible via the hydrogen transfer. The non-canonical DNA base pairs have analogous mechanisms for decay, which have

not been considered in previous studies. These new findings might be important for the interpretation of the experimental data and the design of new experiments aimed to further understand the role of hydrogen transfer in the excited-state dynamics of the DNA. The C-based imino-enol tautomer is both photochemically and thermally unstable, which is a desirable property for the preservation of the genetic code, since this species may originate replication breakdowns and consequently interrupt the proper cell function. Meanwhile, the formation of the highly stable C-based imino-keto tautomer may induce GC $\rightarrow$ AT transitions in the replication process with relevant repercussions in evolutionary biology.

Chapter 8

Resumen

La presente tesis se enmarca dentro de las líneas de investigación del grupo QCEXVAL. Varios temas de diversa índole han sido abordados: metodología, fotofísica de hidruros de boro con aplicación en nanotecnología, y fotoquímica del ADN. En el prefacio se destaca la importancia y el interés de estos estudios. A continuación, se van a resumir los cuatro bloques en que se ha dividido el trabajo presentado en esta tesis: el método RASPT2 (Multiconfigurational Second-Order Perturbation Theory Restricted Active Space) para el cálculo de estados excitados, el método extendido de Hubbard aplicado a hidrocarburos conjugados y aromáticos, la fotofísica de los hidruros de boro *anti-* y *sin-* $B_{18}H_{22}$, y finalmente, la fotoquímica del ADN en procesos de transferencia de protón/hidrógeno en los pares de bases Adenina-Timina y Guanina-Citosina.

El método RASPT2

Debido a la restricción del método CASSCF (Complete Active Space Self-Consistent Field) asociado al limitado número de funciones de estado configu-

racionales adaptadas a la simetría espacial y de spin que comprenden el espacio multirreferencial, es conveniente recurrir a un método más extensivo donde se amplíe el espacio activo, es decir, el número de electrones y orbitales que se requieren en la construcción de las funciones de estado configuracionales. Para ampliar el espacio activo y, por tanto, la aplicabilidad del método, una posibilidad consiste en introducir restricciones en la expansión de la interacción de configuraciones (CI) de la función de onda. Este procedimiento se conoce como RASSCF (Restricted Active Space Self-Consistent Field). El espacio activo se subdivide en tres subespacios: RAS1 que comprende orbitales topológicamente equivalentes a los canónicos doblemente ocupados, RAS2 para orbitales cuya ocupación toma valores entre dos y cero, y RAS3 para orbitales topológicamente equivalentes a los canónicos desocupados. Los orbitales en RAS2 son tratados como en los cálculos CASSCF, es decir, el procedimiento full CI se aplica sobre ellos. Las configuraciones obtenidas de esta manera son las primeras que son esenciales para describir los efectos de correlación no dinámica. Por otra parte, los orbitales en RAS1 y RAS3 son tratados por configuraciones excitadas: un número máximo de excitaciones, comúnmente hasta cuatro, pueden ocurrir desde RAS1 a los otros dos subespacios, y también, un número máximo de excitaciones, hasta cuatro, pueden tener lugar en RAS3 desde RAS1 y RAS2. Ver Fig. 1.1. El principio variacional se utiliza para optimizar tanto los coeficientes de la expansión de la interacción de configuraciones como los orbitales moleculares de la función de onda RAS.

Dado que el número máximo de configuraciones excitadas en RAS1 y RAS3 es limitado, el número de funciones de estado configuracionales puede ser menor que en el método CASSCF sobre el mismo espacio activo. Esta característica permite cálculos multiconfiguracionales con espacios activos más grandes y, por tanto, se extiende la aplicabilidad de las aproximaciones multiconfiguracionales a una gran cantidad de nuevos sistemas y problemas químicos. Es por ello que CASSCF se refiere como un caso particular de RASSCF.

Por ahora, como la correlación dinámica no se recupera en la función de onda RAS, la utilidad del método RASSCF se limita al cálculo de optimizaciones

geométricas o para seleccionar un espacio activo basado en la ocupación de los orbitales naturales para cálculos CASSCF. La incorporación del tratamiento perturbativo a segundo orden a la formulación RASSCF, que origina el método RASPT2, amplia su aplicabilidad a problemas químicos cuantitativos, ya que se tiene en cuenta también la correlación dinámica.

El modelo RASPT2 extiende la teoría perturbativa de muchos cuerpos hasta segundo orden de precisión. Esta teoría tiene la característica de ser de tamaño extensivo, es decir, escala correctamente con el número de partículas del sistema. Se ha probado que el método CASPT2 es de tamaño extensivo, a diferencia de métodos truncados en la expansión de la configuración de interacciones, como MRCI (multireference CI). Debido a las restricciones en el nivel de excitación en RAS1 y RAS3, los problemas de tamaño extensivo pueden aparecer en RASPT2 a menos que se incluyan todas las excitaciones internas.

En la teoría perturbativa de muchos cuerpos, el Hamiltoniano viene dado por el Hamiltoniano de orden cero al que se le añade una pequeña perturbación. En la corrección de la energía a segundo orden aparecen la función de onda de orden cero y la función de onda de primer orden, que, respectivamente, corresponden en el método CASPT2, a la función de onda de referencia CAS y a la función de onda obtenida a partir de todas las configuraciones que interaccionan directamente con la función de referencia CAS. En el método RASPT2 se procede de forma similar.

En ciertos casos, es necesario aplicar algunas correcciones sobre el Hamiltoniano de orden cero debido a la presencia de estados intrusos en el cálculo de estados excitados y también debido a una descripción desbalanceada del Hamiltoniano de orden cero para estructuras electrónicas de capa cerrada con respecto a las de capa abierta.

Existen dos casos de estados intrusos, fuertes y débiles. Los estados intrusos fuertes se caracterizan por una gran interacción con la función de referencia

y, por tanto, la contribución a la energía de segundo orden no es despreciable. La única solución consiste en incrementar el espacio activo incluyendo los orbitales que causan los pequeños denominadores en la energía de segundo orden. Si esto no se puede hacer debido al límite en el tamaño del espacio activo, la teoría perturbativa a segundo orden no se puede utilizar para tratar el problema. Los estados intrusos débiles son más comunes. Su interacción con la función de referencia es pequeña. Es posible eliminar los estados intrusos por una técnica de level shift, que consiste en añadir una constante, ε, al Hamiltoniano de orden cero resultando la energía de segundo orden tal como expresa la ecuación 1.14.

Como se ha mencionado anteriormente, también existe un error sistemático en el método derivado de la descripción desbalanceada del Hamiltoniano de orden cero para estructuras electrónicas de capa cerrada en comparación con estructuras que tienen electrones desapareados. La energía de correlación se sobrestima en los casos de capa abierta y se corrige introduciendo el parámetro IPEA, cuyo valor óptimo es 0.25 ua.

Una última consideración a tener en cuenta en el método RASPT2 es la no ortogonalidad de los estados de la misma simetría espacial y de spin resultantes de este procedimiento partiendo de una referencia RASSCF. Como consecuencia, las soluciones obtenidas ignoran el acoplamiento entre estados. La técnica Multi-state (MS) representa una extensión del método RASPT2 (CASPT2) para el tratamiento perturbativo de situaciones químicas que necesitan dos o más estados de referencia. Por ejemplo, situaciones como los cruces evitados, o la casi degeneración de estados de valencia y Rydberg. Sin embargo, hay que tener precaución al utilizarlo en el cálculo de puntos de cruce entre dos superficies, como en el caso de intersecciones cónicas, crucial en fotoquímica.

Los estados excitados electrónicos de un número de moléculas orgánicas (porfina, eteno, benceno) e inorgánicas (el átomo de níquel y el dianión de tetracloruro de cobre) han sido calculados a nivel RASSCF/RASPT2/MS-RASPT2 utilizando diferentes particiones del espacio activo y distintas estrategias. Los

resultados han sido comparados con valores obtenidos con procedimientos multiconfiguracionales bien establecidos como CASSCF/CASPT2/MS-CASPT2 o monorreferenciales como CCSD, así como con valores experimentales.

La porfina se ha estudiado llevando a cabo diferentes procesos de partición del espacio activo. RASPT2 es una excelente estrategia para evitar divisiones arbitrarias del espacio π en un sistema en el que todo el espacio π (26 electrones en 24 orbitales moleculares (OMs)) está fuera del alcance del método CASPT2. Se ha probado que una adecuada definición del espacio RAS2 es crucial para conseguir precisión en los cálculos. Cuando se calculan diferencias de energía a nivel RASPT2, la mayor precisión se obtiene cuando los OMs que más cambian su número de ocupación de un estado (típicamente el estado fundamental) a otro (un estado excitado) se incluyen todos ellos en el espacio RAS2, dejando los restantes OMs menos significativos en los espacios RAS1/RAS3. Si esto se cumple, el nivel simple-doble (SD) de excitaciones en estos dos últimos espacios es suficiente para conseguir una alta precisión. En la porfina, el espacio de Gouterman constituido por los OMs HOMO, HOMO-1, LUMO, y LUMO+1 forman el conjunto básico para describir los cuatro estados excitados $\pi\pi^*$ más bajos en energía, y por tanto, será suficiente incluirlos en RAS2 mientras se dejan los restantes OMs $\pi\pi^*$ en RAS1/RAS3 y permitiendo hasta dobles excitaciones para obtener resultados precisos. Sin embargo, la descripción de estados más altos en energía requieren una extensión del espacio RAS2 para incluir OMs adicionales. Si los OMs adecuados se excluyen de RAS2 y el nivel SD de excitaciones es el permitido en RAS1/RAS3, entonces, se obtiene un tratamiento CI desbalanceado y la corrección perturbativa a segundo orden no es capaz de compensar los resultados. Sin embargo, es posible mejorar los resultados cuando se utilizan diferentes niveles de excitación para los dos estados considerados, como SD para el estado fundamental y SDT para el estado excitado, o aumentar el nivel de excitaciones en RAS1/RAS3 lo más que se pueda, SDT o incluso SDTQ, aunque esta última estrategia conlleva espacios configuracionales muy grandes y, por tanto, es imposible su aplicación. Estos resultados abren la posibilidad de utilizar RASPT2 para muchos sistemas orgánicos con extensivos espacios π sin pérdida de precisión.

Cálculos de estados excitados singlete y triplete de valencia y Rydberg se han llevado a cabo en el eteno y el benceno con el método RASPT2. Para ello se ha utilizado una nueva estrategia de partición del espacio RAS en el que los OMs Rydberg, típicamente nueve ($n = 3$) para sistemas orgánicos comunes, se sitúan en RAS3, dejando en RAS2 el espacio de OMs y electrones de valencia $\pi\pi^*$, y permitiendo, a parte del full CI en RAS2, solo simples excitaciones hacia RAS3. Los orbitales Rydberg pueden situarse fuera del espacio RAS2 y, por tanto, el esfuerzo computacional decrece sustancialmente. Los cálculos son más simples cuando se comparan con sus análogos CASPT2 porque en RASPT2 se utiliza un único espacio activo para las distintas simetrías. Sin embargo, esta aproximación no resuelve los problemas de mezcla valencia-Rydberg ya presentes en CASSCF/CASPT2, dando altas energías de excitación y funciones de onda mezcladas con extensiones orbitales grandes para algunos estados de valencia. Cuando solo los OMs $\pi\pi^*$ se incluyen en el espacio RAS2, el procedimiento multi-state, MS-RASPT2, es necesario para resolver la mezcla y proporcionar estados ortogonales claramente de valencia o Rydberg. En el caso del eteno, la inclusión de los OMs $\sigma\sigma^*$ en RAS1 y RAS3 determina que RASPT2(SD) no es suficiente para describir las excitaciones $\sigma\sigma^*$. Aumentando el nivel de excitación a SDT se resuelve el problema para el caso del eteno, pero no puede decirse que sea una regla general. Cuando se incluye la correlación π y σ en el espacio de referencia CI, la mezcla valencia-Rydberg se resuelve a nivel RASPT2(SDT), sin la necesidad de utilizar el procedimiento multi-state. Esto muestra la importancia de la correlación electrónica en la definición de la función de onda cuando se trata el problema de mezcla valencia-Rydberg.

Con respecto al cálculo de sistemas con metales de transición de la primera fila, el interés se centra en el efecto de las energías de excitación cuando se mueve la capa correlacionada $4d$ de RAS2 a RAS3. Los espectros electrónicos del átomo de níquel y el dianión tetracloruro de cobre han sido analizados. Los cálculos RASPT2 reproducen bien los correspondientes cálculos CASPT2, aunque se observan algunos casos excepcionales, como por ejemplo, el estado $^1\mathrm{S}(3\mathrm{d}^{10})$ del átomo de níquel.

El modelo de Hubbard

En los años cincuenta hubo un gran interés en la teoría de efectos de correlación en el gas de electrones libres, utilizado como modelo para las bandas de conducción de metales y aleaciones. Sin embargo, los metales de transición y tierras raras tienen, aparte de su banda de conducción, las bandas d- y f- parcialmente llenas, y el gas de electrones libres no proporciona un buen modelo para estas bandas. Se requiere, pues, una teoría de correlaciones que tenga en cuenta adecuadamente la naturaleza atomística del sólido. En 1963, John Hubbard propuso un modelo para determinar el balance entre el comportamiento de banda y el atómico en las bandas d de los metales de transición. Básicamente, el modelo de Hubbard establece que los electrones pueden promocionar entre sitios, representados por iones, y que están dispuestos en un patrón ordenado y cristalino. Cuando dos electrones ocupan el mismo sitio, existe una repulsión mutua que introduce correlaciones adicionales entre electrones más allá de las debidas al principio de exclusión de Pauli. Si la repulsión electrón-electrón es débil, entonces los electrones vagan más o menos libres alrededor del material que se comporta como un metal. Pero si la repulsión es fuerte, los electrones se localizan forzosamente en posiciones atómicas fijas y el material se comporta como un aislante. En un sistema *mitad lleno* (un electrón por sitio), el modelo predice la formación de un aislante de Mott, como algunos óxidos de metales de transición. Cuando se dopa un aislante de Mott, bien añadiendo o eliminando un pequeño número de electrones, éste pasa a tener un comportamiento metálico.

Para estudiar la dinámica de los electrones de conducción, se define el Hamiltoniano de la ecuación 1.29 que describe el problema de un gas de electrones interactuante sumido en un sistema de estado sólido. Cuando la interacción electrón-electrón no se tiene en cuenta, las funciones de Bloch resultan ser las funciones propias del Hamiltoniano electrónico. Hay dos clases de materiales donde las funciones de Bloch se modifican cuando la repulsión electrón-electrón se añade al Hamiltoniano: sistemas de electrones casi libres y sistemas fuerte-

mente enlazados. El interés de la presente tesis se centra en el segundo grupo de materiales.

Suponiendo que los electrones están fuertemente enlazados a los centros de la red, es conveniente expandir la función de onda que describe el sistema en términos de los estados de Wannier que, a diferencia de las funciones de Bloch, no son funciones propias del Hamiltoniano electrónico. Cuando el solapamiento interatómico es despreciable, es decir, el potencial se aproxima a una superposición de potenciales atómicos independientes, la función de Wannier $\psi_{\mathbf{R}n}(\mathbf{r}) \equiv \langle \mathbf{r}|\psi_{\mathbf{R}n}\rangle$ es igual al enésimo orbital de un átomo aislado centrado en la coordenada $\mathbf{R}$. Sin embargo, para un acoplamiento interatómico finito, los N estados degenerados etiquetados por n se desacoplan para formar una banda de energía (Fig. 1.2). Cuando la energía de Fermi se sitúa entre dos bandas energéticamente separadas, el material es un aislante. Pero cuando la energía de Fermi se localiza dentro de una banda, el sistema es metálico.

El Hamiltoniano de la ecuación 1.38 define la representación de sistemas fuertemente enlazantes. Por otra parte, la expresión 1.42 define el Hamiltoniano de Hubbard en el límite atómico, donde los átomos están bien separados y el solapamiento de orbitales vecinos es débil.

El Hamiltoniano de Hubbard no sólo se utiliza en física del estado sólido sino también en química como modelo de hidrocarburos π-conjugados. En la presente tesis, se hace uso del modelo extendido de Hubbard, donde se añade al Hamiltoniano de la ecuación 1.42 un término de repulsión electrónica para dos electrones en sitios vecinos, dando lugar a la expresión 1.43. Este Hamiltoniano tiene una solución analítica para sistemas con dos sitios idénticos y dos electrones, dando como resultado los valores propios de energía de los tres estados energéticamente más bajos; estado fundamental, y estados excitados triplete y singlete. La molécula de eteno es un sistema π-conjugado con dos sitios y dos electrones activos en el modelo de Hubbard y, por tanto, es un buen candidato para, a partir de las superficies de energía potencial de los tres estados mencionados anteriormente, estimar los parámetros del Hamiltoniano

de Hubbard y extender su aplicabilidad al cálculo de estructuras y energías de excitación de otros compuestos π-conjugados.

Una simple extensión del modelo de Hückel, añadiendo la repulsión electrón-electrón como se describe en el Hamiltoniano de Hubbard de la expresión 1.42, da lugar a un atractivo modelo del enlace químico. La forma extendida del modelo de Hubbard, incluyendo también las interacciones electrón-electrón de vecinos más próximos, ecuación 1.43, puede ser parametrizada a partir de curvas de energía potencial precisas para un fragmento de dos sitios. Las curvas de energía potencial para el eteno (Fig. 6.2) han sido calculadas utilizando el método multiconfiguracional MS-RASPT2, esencial para tratar adecuadamente la mezcla valencia-Rydberg en los estados $\pi\pi^*$ del eteno y de otras pequeñas moléculas. Los parámetros del modelo de Hubbard, extraídos de las curvas de energía potencial del eteno en función de la longitud de enlace carbono-carbono, podrían aplicarse a hidrocarburos arómaticos y conjugados así como a especies de carbono como el grafeno.

En cálculos del butadieno, hexatrieno y benceno, el modelo extendido de Hubbard ofrece una descripción razonable de las longitudes de enlace carbono-carbono, un valor intermedio entre las longitudes de enlace a nivel MP2 y las longitudes de enlace experimentales. El modelo de Hubbard también proporciona una muy buena descripción de las energías de excitación más bajas cuando se comparan con cálculos MS-RASPT2 y valores experimentales. Sin embargo, el modelo no reproduce correctamente el estado ^1B_u en el butadieno y el hexatrieno, ilustrando una limitación de los modelos de vecinos más próximos. El modelo de Hubbard es sólo fiable para estados covalentes, es decir, estados neutros descritos por configuraciones doblemente excitadas.

A pesar de la limitación en la descripción de los estados iónicos en polienos, el modelo extendido de Hubbard incluye satisfactoriamente las interacciones más importantes necesarias para entender el enlace químico. Este modelo es capaz de describir simultáneamente la estructura geométrica y las excitaciones electrónicas de más baja energía. El éxito del modelo se debe a la inclusión

de la correlación electrónica en el Hamiltoniano efectivo. Para pequeñas moléculas, como las consideradas en el presente trabajo, son posibles soluciones exactas, y para sistemas de mayor tamaño molecular el coste computacional es menor comparado con un cálculo *ab initio*.

Hidruros de Boro y la fotofísica de los isómeros $B_{18}H_{22}$

Entre 1912 y 1936, Alfred Stock y sus colaboradores prepararon una serie de compuestos de boro, conocidos como boranos o hidruros de boro, que se caracterizan por ser muy reactivos frente al oxígeno molecular. Diseñaron experimentos en vacío para manipular hidruros de boro sin presencia de oxígeno ni humedad en los que caracterizaron los siguientes boranos: B_2H_6, B_4H_{10}, B_5H_9, B_5H_{11}, B_6H_{10} and $B_{10}H_{14}$, que se muestran en la Fig. 1.3. Cada átomo de boro está enlazado a, al menos, un átomo de hidrógeno y, a veces, estos átomos de hidrógeno establecen puentes de hidrógeno entre dos átomos de boro. La reactividad de los boranos es muy diversa, algunas reacciones son: la formación de aniones de boro, la sustitución de átomos de hidrógeno por halógenos, la formación de aductos o la adición de heteroátomos (C, Si, P, etc.) en el esqueleto de los boranos.

Con respecto a las características estructurales y de enlace, los clústeres de boro son sistemas moleculares compuestos de átomos de boro situados en los vértices de poliedros de cara triangular (deltaedros). Cada átomo de boro contribuye con un número máximo de tres electrones a sus vecinos más próximos, que pueden ser hasta cinco, seis o incluso siete átomos. Por tanto, las conexiones entre átomos no se pueden considerar enlaces simples de un par de electrones. Este problema se conoce como deficiencia electrónica y se resuelve con el concepto de enlace multicentro (ver sección 1.3).

Los clústeres simples de boranos son, o bien deltaedros esféricos (*closo*-boranos) o bien resultan de eliminar uno, dos o tres vértices de alta conectividad de las estructuras *closo*- para dar estructuras *nido*-, *arachno*- e *hypho*-, respectiva-

mente. En la Fig. 1.4 estas estructuras están conectadas mediante flechas diagonales. Existe una relación entre la geometría de los hidruros de boro y su estructura electrónica establecida por la regla de Wade. Para un borano con p vértices, el número de pares de electrones del esqueleto depende de la forma del clúster, como se puede ver en la Tabla 1.1.

A partir de clústeres pequeños se pueden obtener clústeres más grandes que reciben el nombre de macropoliedros. Un número importante de boranos macropoliédricos, conocidos con la fórmula general B_nH_{n+4}, consisten en dos unidades *nido* que comparten dos vértices. Experimentalmente, el mayor clúster *nido* conocido es $B_{11}H_{15}$ mientras que el más pequeño borano *nido:nido*-macropoliédrico es $B_{12}H_{16}$, que se muestran en la Fig. 1.5.

Cualquier borano macropoliédrico *nido:nido* neutro es menos estable que su correspondiente simple clúster *nido*-borano, siempre y cuando el número total de vértices sea igual o menor que 11. Para 12 o más vértices, hasta 19, los boranos *nido:nido* son energéticamente favorables frente a sus correspondientes clústeres simples. Los puentes de hidrógeno en las caras abiertas de los clústeres de boro prefieren enlazar vértices de menor conectividad. A continuación, se introducirán las características más significativas del hidruro de boro estudiado en la presente tesis, *nido*(10):*nido*(10)-octadecaborano(22) (Fig. 1.6).

A principios de los sesenta, los dos isómeros $B_{18}H_{22}$ fueron preparados y caracterizados por estudios de difracción de rayos X. Estos dos hidruros de boro están estructuralmente relacionados con el *nido*-decaborano(14) (Fig. 1.7) y se pueden generar por la fusión de dos fragmentos decaboranos en la que se comparten dos átomos (5-6) como muestra la Fig. 1.6. El isómero *anti* es centrosimétrico mientras que la especie *sin* tiene simetría C_2.

Una característica estructural de *anti*- y *sin*-$B_{18}H_{22}$ es la coordinación que tienen los dos átomos de B (5-6) del vértice compartido con otros seis átomos de B. En la Fig. 1.8 se puede observar este hecho.

La química de ambos isómeros *anti*- y *sin*-$B_{18}H_{22}$ es diversa, tal como han estudiado Hawthorne y sus colaboradores. Por ejemplo, con respecto a su carácter ácido-base, la valoración de ambos isómeros en medio acuoso indica que se comportan como ácidos dipróticos. Por otra parte, a diferencia del decaborano(14), en el octadecaborano(22) no se tiene evidencia de formación de derivados en presencia de agentes nucleofílicos como el acetonitrilo. También se ha observado que en reacciones de reducción de *anti*- y *sin*-$B_{18}H_{22}$ con amalgama sódico en tetrahidrofurano se producen iones $[B_{18}H_{22}]^{2-}$.

Una de las aplicaciones más comunes de $B_{18}H_{22}$ está asociada a la fabricación de semiconductores. Hidruros de boro de gran tamaño, al menos un clúster de cinco átomos de boro, pueden ser materias primas para impurezas en semiconductores de silicio de tipo p.

El trabajo presentado en esta tesis que combina conjuntamente el estudio experimental y computacional sobre la fotofísica de los hidruros de boro $B_{18}H_{22}$ racionaliza la observación de que el isómero *anti* presenta una intensa fluorescencia, mientras que el isómero *sin* no muestra tal actividad (Fig. 6.12). El tiempo de vida media medido para el estado excitado singlete más bajo en *anti*-$B_{18}H_{22}$ es dos órdenes de magnitud mayor que en el correspondiente isómero *sin*, y este último decae no radiativamente en la escala de tiempo del picosegundo (75 ps). Cálculos químico-cuánticos basados en la metodología CASSCF/CASPT2 indican que las diferencias en las geometrías de los dos isómeros en los puntos de intersección cónica calculados a lo largo de las hipersuperficies de energía potencial de S_0 y S_1 es el factor determinante en los diferentes procesos de decaimiento. Así pues, el isómero *sin* muestra una gran elongación B-B de alrededor 1.2 Å en la intersección cónica, mientras que el isómero *anti* presenta elongaciones B-B menores de 1 Å en su correspondiente intersección cónica (Fig. 6.13). La energía absorbida en S_1 en *anti*-$B_{18}H_{22}$ es, por tanto, redistribuida vibracionalmente entre varias regiones de la molécula en vez de en un solo modo vibracional, tal como sucede en el caso del isómero *sin*. El tiempo de vida media del estado S_1 en el isómero *anti* es más dilatado que en el isómero *sin*, permitiendo la relajación del primero via fluorescencia.

El interés del presente trabajo está relacionado con la síntesis de derivados de los isómeros $B_{18}H_{22}$ con sustituyentes que no alteran las geometrías de clúster de los hidruros de boro, con el fin de obtener un eficiente acoplamiento spin-órbita entre los estados singlete y triplete que finalmente dan lugar a la producción de oxígeno singlete.

Fotoquímica de los pares de bases de ADN en procesos de transferencia de protón/hidrógeno

La estructura del ADN como doble hélice se mantiene por dos tipos de interacciones, enlaces por puentes de hidrógeno y apilamiento de tipo π. Los enlaces de hidrógeno median la formación de pares de bases entre los ácidos nucleicos Adenina (A) y Timina (T), por un lado, y Guanina (G) y Citosina (C), por otro, en la doble hélice Watson-Crick de ADN. En la Fig. 1.9 se muestran los pares de bases canónicos Watson-Crick de ADN, donde se indican los dos puentes de hidrógeno en el caso de AT y los tres enlaces de hidrógeno en GC.

Los patrones de enlaces de hidrógeno alterados o pares de bases que no son Watson-Crick pueden dar lugar a la modificación del código genético provocando mutaciones. Un par de ejemplos de enlaces de hidrógeno alterados se muestran en la Fig. 1.10. Estos patrones corresponden a la formación de pares de tautómeros raros de ADN que, según se piensa, causan la formación de emparejamientos defectuosos que producen mutaciones espontáneas en el código genético.

A comienzos de 1963, P.-O. Löwdin introdujo la hipótesis de que la doble transferencia de protón en pares de bases de ADN podía ser una posible fuente de mutaciones espontáneas. Desde entonces se han llevado a cabo estudios teóricos para demostrar la hipótesis de Löwdin, que se mencionan en la introducción (sección 1.4). Todos los estudios coincidían en que la transferencia de un solo protón era menos favorable que la doble transferencia de protón.

Sin embargo, la barrera energética de la doble transferencia de protón seguía siendo alta.

En los noventa, el proceso de doble transferencia de protón por fotoexcitación en pares de bases de ADN y en sistemas modelo como el 7-azaindole fue de gran interés, tanto experimentalmente como teóricamente. El tema central de la discusión de estos trabajos ha sido dilucidar si el mecanismo de esta reacción es por pasos o concertado. De hecho, en un trabajo sobre el 7-azaindole de Serrano-Andrés et al. se encontraron dos mecanismos estáticos por pasos basados en la presencia de una especie intermedia neutra y otra iónica después que un protón se transfiriera de una parte de la molécula a otra. En un estudio teórico de Guallar et al. sobre los pares de bases GC y AT, se encontró un estado electrónico de transferencia de carga a bajas energías en la región correspondiente a la transferencia de un solo protón que neutraliza la separación de carga inducida por la transición electrónica. Domcke y sus colaboradores caracterizaron una intersección cónica en GC y AT entre el estado de transferencia de carga y el estado fundamental después de la transferencia de un solo protón. Así pues, la doble transferencia de protón ocurre, precisamente, para recuperar el estado fundamental de los pares de bases canónicos cuando son fotoinducidos. En un estudio del sistema piridina-pirrol, Frutos et al. encontraron un cruce evitado entre el primer estado excitado y el estado de transferencia de carga, a nivel CASSCF/CASPT2, que conecta ambos estados no adiabáticamente. Estos trabajos teóricos mencionados explican la fotoestabilidad de los pares de bases Watson-Crick, y han sido respaldados por trabajos experimentales usando técnicas de medición en la escala de tiempo del femtosegundo (sección 1.4). Sin embargo, actualmente, está en tela de juicio el papel de los enlaces por puentes de hidrógeno en la dinámica de la rápida desactivación del ADN fotoinducido. Brevemente, algunos experimentos señalan que la transferencia de protón o hidrógeno no es un proceso relevante en la fotoestabilidad del ADN, siendo otros mecanismos de desactivación más favorables. Además, el efecto del medio pone de manifiesto la importancia del entorno (ADN o solución acuosa) en la forma en que los procesos de desactivación encontrados en vacío son modificados.

Con respecto a la malformación de pares de tautómeros en el par de bases GC, Florián y Leszczynsky determinaron, a nivel HF/MP2, que los pares de bases más bajos en energía, entre los que resultan de diferentes disposiciones de los protones compartidos en los enlaces de hidrógeno, son el par de bases canónico GC y el tautómero raro imino-keto/amino-enol G^*C^*. Dos caminos energéticos diferentes pueden gobernar las disposiciones de los protones para conectar el par canónico GC y el par de bases G^*C^*. El camino de más alta energía viene mediado por un par de bases iónico GC_{ionic}, y tiene dos estados de transición que conectan GC con GC_{ionic} y GC_{ionic} con G^*C^*, como muestra la Fig. 1.11. El camino de más baja energía es un proceso concertado que conecta GC y G^*C^* mediante un estado de transición, como también muestra la Fig. 1.11.

Kryachko y Sabin propusieron un mecanismo para explicar el origen de la malformación de pares en AT, el cual consiste en cuatro fases. Las especies AT y A^*T^* están conectadas por un intermedio iónico AT_{ionic} en un proceso compuesto por dos pasos que constituyen las dos primeras fases del mecanismo. Sin embargo, a nivel HF la especie A^*T^* es más inestable energéticamente que el intermedio AT_{ionic}. La presencia de un estado de transición cercano a AT_{ionic} afecta fuertemente a su metaestabilidad. A partir de este estado de transición se activa el mecanismo de valle de la malformación de pares, el cual cubre las tercera y cuarta fases del proceso de tautomerización de AT. Se produce un desplazamiento en el plano molecular de las bases A y T de manera que los puentes de hidrógeno se establecen entre el grupo amino de la A y el nitrógeno de la T, y el grupo imino de la A y un oxígeno de la T, dando lugar a la especie AT_3 como muestra la Fig. 1.13. Así se completa la tercera fase. A partir de esta especie, por transferencia de protón, pueden producirse las especies AT_4 y AT_5, dependiendo del protón que se haya transferido. Como AT_4 es más estable que AT_5, el sistema finalizará el proceso (cuarta fase) con la tautomerización AT_4.

Los mecanismos de desactivación no radiativa en las nucleobases aisladas y la formación de lesiones en dímeros de pirimidinas con apilamiento de tipo π

explican, junto con los resultados sobre los pares de bases de ADN presentados en esta tesis, la fotoquímica y la fotorreactividad del ADN.

En el presente trabajo, se establecen tres posibles mecanismos para el decaimiento de la energía en los pares de bases de ADN: 1) doble transferencia de protón por pasos (SDPT), 2) doble transferencia de hidrógeno por pasos (SDHT), 3) doble transferencia concertada de protón (CDPT). Según los resultados hallados, el mecanismo SDHT, es el más favorable para explicar las propiedades fotoestables de los pares de bases canónicos de ADN y, también, la formación de tautómeros. En el mecanismo SDHT, la transferencia de protones desde las purinas a las pirimidinas favorece la evolución de los estados excitados energéticamente más bajos de carácter localmente excitado a transferencia de carga, que posteriormente conduce a una transferencia de hidrógeno para formar los correspondientes intermedios neutros (Fig. 6.25 - 6.27). Estas especies intermedias se localizan en una región de casi degeneración con la superficie del estado fundamental, lo que puede canalizar el camino de reacción hacia el par de bases canónico de ADN o hacia una nueva forma de tautómero, a través de la transferencia de un segundo hidrógeno en el mismo o diferente enlace de hidrógeno entre bases, respectivamente. De forma análoga, los procesos SDHT a partir de las formas tautoméricas explican la formación de los pares de bases canónicos o el restablecimiento de los tautómeros. Cabe señalar que la especie imino-enol con respecto a C en la Fig. 6.27 representa una configuración alta en energía entre los pares de bases GC estudiados, con la barrera energética más baja para restablecer la forma Watson-Crick. Según los perfiles energéticos obtenidos (Fig. 6.20 - 6.30), las conversiones térmicas son más favorables a través del mecanismo SDPT que con los mecanismos SDHT y CDPT.

En solución con solventes polares, según estudios previos, el mecanismo SDHT en GC parece inactivar el canal de conversión interna que conduce el par de bases GC a su estructura canónica o a sus nuevas configuraciones tautoméricas. Sin embargo, el dímero GC incrustado en el ADN, estudiado con la aproximación QM/MM, manifiesta un comportamiento similar con respecto al par de bases aislado para los caminos fotoquímicos de decaimiento energético que

implican doble transferencia de protón/hidrógeno, respaldando observaciones experimentales y trabajos teóricos previos. Los fenómenos de transferencia de hidrógeno en los pares de bases GC parecen ser más favorables a lo largo del enlace de hidrógeno central. Los tiempos de vida media para la transferencia de hidrógeno (sobre 50 fs) son más cortos que los decaimientos no radiativos localizados en los monómeros. Sin embargo, una comparación directa de tiempos de vida media en el estado excitado medidos experimentalmente no es posible, ya que estos pueden verse afectados por procesos fotoquímicos intracadena, como la formación de excipletes.

En conclusión, nuevas rutas fotoquímicas son accesibles en los pares de bases de ADN cuando incide radiación ultravioleta, que competirán con los caminos de decaimiento energético de las nucleobases aisladas y de dímeros con apilamiento de tipo π. Los procesos estudiados en la presente tesis serán más eficientes en conformaciones donde los enlaces de hidrógeno son importantes en la doble cadena de ADN. Vibraciones en los enlaces de hidrógeno pueden activar las transferencias de protón/hidrógeno en los pares de bases de ADN, tal como sugieren las simulaciones dinámicas del dímero GC (Tabla 6.15). Entre los mecanismos estudiados, el más favorable es el de doble transferencia de hidrógeno que proporciona al sistema nuevas rutas fotoestables, además de las presentes en las nucleobases aisladas. Pero la formación de tautómeros, como ocurre en vacío, también es posible a través de la transferencia de hidrógeno. Los pares de bases no canónicos tienen análogos mecanismos de decaimiento, que no han sido previamente considerados. Estos nuevos resultados puede que sean importantes para la interpretación de valores experimentales y en el diseño de nuevos experimentos dirigidos a entender el papel de la transferencia de hidrógeno en la dinámica del ADN en el estado excitado. El tautómero imino-enol con respecto a C es, tanto fotoquímicamente como térmicamente, inestable, lo que es beneficioso para la preservación del material genético, ya que estas especies pueden interrumpir la replicación. Por otra parte, la formación del tautómero estable imino-keto con respecto a C puede inducir transiciones GC $\rightarrow$ AT en el proceso de replicación, lo que puede tener importantes repercusiones en la biología evolutiva.

List of Terms

A

A	Adenine / Absorbance
ad	adiabatic
AT	Adenine-Thymine
AC	Avoided Crossing
AMFI	Atomic Mean-Field Integral
ANO	Atomic Natural Orbital
ANO-L	Large Atomic Natural Orbital
ANO-RCC	Relativistic Atomic Natural Orbital
ANO-S	Small Atomic Natural Orbital
aug-cc-pVDZ	augmented correlation consistent polarized Valence Double Zeta

B

B3LYP	Becke 3-parameter Lee-Yang-Parr functional

C

C	Cytosine
CAS	Complete Active Space
CAS CI	Complete Active Space Configuration Interaction
CASSCF	Complete Active Space Self-Consistent Field
CASSI	Complete Active Space States Interaction

CASPT2	Second-Order Perturbation Theory Complete Active Space
cc	correlation consistent
CC	Coupled-Cluster
cc-pVXZ	correlation consistent polarized Valence X Zeta, where X takes the labels: D (double), T (triple), Q (quadruple), 5 (quintuple), 6 (sextuple)
cc-pCVXZ	correlation consistent polarized Core-Valence X Zeta, where X takes the labels: D (double), T (triple), Q (quadruple), 5 (quintuple), 6 (sextuple)
CCSD	Coupled-Cluster Singles and Doubles
CCSD(T)	Coupled-Cluster Singles, Doubles, and Non-iterative Triples correction
CCSDT	Coupled-Cluster Singles, Doubles, and Triples
CCSDTQ	Coupled-Cluster Singles, Doubles, Triples, and Quadruples
CDPT	Concerted Double Proton Transfer
CGTO	Contracted Gaussian Type Orbital
CI	Configuration Interaction / Conical Intersection
CID	Configuration Interaction Doubles
CIS	Configuration Interaction Singles
CISD	Configuration Interaction Singles and Doubles
CISDT	Configuration Interaction Singles, Doubles, and Triples
CPD	Cyclobutane Pyrimidine Dimer
CSF	Configuration State Function
CT	Charge Transfer

D

D	Double
DFT	Density Functional Theory
DH	Dunning-Huzinaga
DNA	Deoxyribose Nucleic Acid
DPT	Double Proton Transfer
DZ	Double Zeta

DZP	Double Zeta plus Polarization

E

EA	Electron Affinity
EE	Electrostatic Embedding
EET	Electronic Energy Transfer
E_F	Fermi Energy

F

F	Fluorescence
FC	Franck-Condon
FF	Force Field

G

G	Guanine
GC	Guanine-Cytosine
GHO	Generalized Hybrid Orbital
GS	Ground State
GTO	Gaussian Type Orbital

H

HL	Hydrogen Link
HF	Hartree-Fock
HMO	Hückel Molecular Orbital
HOMO	Highest Occupied Molecular Orbital

I

IC	Internal Conversion
INT	Intermediate
IP	Ionization Potential
IPEA	Ionization Potential Electron Affinity
IR	Infrared
ISC	Intersystem Crossing
IVR	Intramolecular Vibrational Relaxation

L

LE	Locally Excited
LF	Ligand Field
LIIC	Linear Interpolation in Internal Coordinates
LSCF	Local Self-Consistent Field
LA	Link Atom
LCAO	Linear Combination of Atomic Orbitals
LCAO-MO	Linear Combination of Atomic Orbitals-Molecular Orbitals
LS	Level Shift
LUMO	Lowest Unoccupied Molecular Orbital

M

MBPT	Many-Body Perturbation Theory
MBS	Minimum Basis Set
MCSCF	Multi-Configurational Self-Consistent Field
MD	Molecular Dynamics
ME	Mechanical Embedding
MECP	Minimum Energy Crossing Point
MM	Molecular Mechanics
MO	Molecular Orbital
MP2	Møller-Plesset perturbation theory to Second Order
MP4	Møller-Plesset perturbation theory to Fourth Order
MRCI	Multi-Reference Configuration Interaction
MR-MBPT	Multi-Reference Many-Body Perturbation Theory
MS	Multi-State
MS-CASPT2	Multi-State Second-Order Perturbation Theory Complete Active Space
MS-RASPT2	Multi-State Second-Order Perturbation Theory Restricted Active Space
MXS	Minimal in the Crossing Seam

N

NAB	Nucleic Acid Base

NIST	National Institute of Standards and Technology

P

PDT	Photodynamic Therapy
PGTO	Primitive Gaussian Type Orbital
PE	Polarized Embedding
PEH	Potential Energy Hypersurface
PES	Potential Energy Surface
PMCAS	Perturbed Modified Complete Active Space

Q

Q	Quadruple
QCEXVAL	Quantum Chemistry of the Excited State University of Valencia
QDPT	Quasi-Degenerate Perturbation Theory
QM	Quantum Mechanics
QZ	Quadruple Zeta

R

rad	radiative
RAS	Restricted Active Space
RASSCF	Restricted Active Space Self-Consistent Field
RASPT2	Second-Order Perturbation Theory Restricted Active Space
Ryd	Rydberg

S

S	Single / Singlet
SCE	Self-Consistent Embedding
SCF	Self-Consistent Field
SH	Surface Hopping
SOC	Spin-Orbit Coupling
SDHT	Stepwise Double Hydrogen Transfer
SDPT	Stepwise Double Proton Transfer

SPT	Single Proton Transfer
SS	Single-State
STO	Slater Type Orbital

T

T	Thymine / Triple / Triplet
TAU	Tautomer
TD-DFT	Time-Dependent Density Functional Theory
TDM	Transition Dipole Moment
TET	Triplet-triplet Energy Transfer
TZ	Triple Zeta

U

UHF	Unrestricted Hartree-Fock
UV	Ultraviolet

V

VIS	Visible
VSEPR	Valence Shell Electron Pair Repulsion
VB	Valence Bond
VET	Vibrational Energy Transfer
VXZ	Valence X Zeta, where X takes the labels: D (double), T (triple), Q (quadruple), 5 (quintuple)

W

WC	Watson Crick

List of Tables

1.1 Relation between single-cluster types and their skeletal electron pairs. 44

5.1 Electromagnetic spectrum regions of chemical interest and associated molecular phenomena. 95

6.1 Excitation energies of the singlet and triplet valence $\pi\pi^*$ states of Free Base Porphin (D_{2h}). 119

6.2 Natural occupation numbers of the most relevant molecular orbitals of the low-lying $\pi\pi^*$ states of Free Base Porphin. 121

6.3 Excitation energies of selected states of Ethene (D_{2h}). 124

6.4 Excitation energies of the singlet and triplet valence $\pi\pi^*$ and $n = 3$ Rydberg states of Ethene. 127

6.5 Excitation energies for the low-lying valence and Rydberg singlet states of Benzene (D_{6h}). 129

6.6 Comparison between CASPT2 and RASPT2 excitation energies of the excited states of the Nickel atom. 131

6.7 Excitation energies of $CuCl_4^{2-}$ computed at the CASPT2(11,11) and RASPT2(11,0,m;0,6,5) levels of calculation compared with the available experimental data. 134

6.8 *Ab initio* derived Hubbard parameters. 143

6.9 Comparison of fits of the t parameter. 145

6.10 Carbon-carbon bond lengths of ethene, *trans*-1,3-butadiene, *trans*-1,3,5-hexatriene, and benzene at the ground-state equilibrium geometry. 148

6.11 Excitation energies of the lowest valence singlet and triplet states of ethene, *trans*-1,3-butadiene, *trans*-1,3,5-hexatriene, and benzene at the ground-state equilibrium geometry. 150

6.12 Photophysical properties of *syn*-$B_{18}H_{22}$ and *anti*-$B_{18}H_{22}$ in air-saturated hexane: absorption spectra (molar absorption coefficients ϵ at band maxima λ_{max}), fluorescence band maximum λ_F, quantum yield of fluorescence Φ_F, lifetime of the excited singlet states τ_F, quantum yield the triplet states Φ_T, lifetime of the triplet states τ_T, quantum yield of singlet oxygen formation Φ_Δ. 160

6.13 Theoretical (CASPT2) absorption and emission spectra of *anti*-$B_{18}H_{22}$. 166

6.14 Theoretical (CASPT2) absorption and emission spectra of *syn*-$B_{18}H_{22}$. 169

6.15 NH bond lengths and scalar projections of the velocities of the H'_{41}, H_1, and H_{21} atoms over the lines defined by the $N'_4H'_{41}\cdots O_6$, $N'_3\cdots H_1N_1$, and $O'_2\cdots H_{21}N_2$ hydrogen bonds, respectively, in the GC base pair. The mechanisms and photochemical decay paths observed in the simulations are indicated. 208

List of Figures

1.1 Partition of the CAS and RAS orbital subspaces. 19

1.2 Energy bands for a solid within the range of weakly to finite interatomic coupling. 35

1.3 Stock's structures of boron hydrides. 43

1.4 Polyhedral structures of single-cluster boranes with diagonal arrows indicating the *closo*-, *nido*- and *arachno*- structures connected by successive removal of BH vertex of high connectivity. Hydrogens are not shown. 45

1.5 The largest experimentally known homonuclear *nido* single cluster, $B_{11}H_{15}$, and the smallest experimentally known homonuclear *nido:nido*-macropolyhedral borane cluster, $B_{12}H_{16}$. 46

1.6 Structure and numbering for the *anti*-$B_{18}H_{22}$ and *syn*-$B_{18}H_{22}$. 47

1.7 Structure and numbering for the $B_{10}H_{14}$. 48

1.8 Lipscomb's structures of *anti*-$B_{18}H_{22}$ and *syn*-$B_{18}H_{22}$. 48

1.9 The canonical AT and GC base pairs of DNA. 51

1.10 The double proton transfers in the canonical AT and GC base pairs of DNA. 52

1.11 The proton transfer PES of the GC base pair at the HF level of theory. 55

1.12 The portion of the PES of the AT base pair at the HF level of theory. 56

1.13 The mispairing valley in the PES of the AT base pair at the HF and DFT/B3LYP levels of theory. 57

3.1 Scheme of the most common quantum-chemical *ab initio* methods. 64

3.2 Scheme of the division of a molecular system into the QM and MM subsystems employed in the QM/MM approach. 72

3.3 Labeling of atoms in the QM and MM regions employed in the QM/MM method when the boundary region breaks a covalent bond and the LA model is used. 74

4.1 Segmented and general contraction scheme for basis sets. . . . 87

5.1 Electromagnetic Spectrum 94

5.2 Main intramolecular deexcitation mechanisms in Photochemistry. 98

5.3 Photochemical processes. 100

5.4 Hyperline resulting from the intersection of two PESs (a) and the energy representation of the excited and ground states against x_1 and x_2 coordinates of the branching space (b). 102

5.5 EET mechanisms preferred according to the distance between the donor and acceptor. 107

5.6 Correlation effects in the CASPT2//CASSCF protocol. 110

6.1 Set of molecules considered in the study of computing excitation energies by using the RASPT2 method. 116

6.2 Potential energy curves of ethene for the three lowest valence states and the lowest Rydberg state performed with RASSCF/MS-RASPT2(12,2,2;4,6,8)(SD) using ANO-L $6s5p4d2f/3s2p1d$ basis set and a set of even-tempered $1s1p1d$ functions on C atoms. 141

6.3 The *ab initio* derived values of the Hubbard t and U^{eff} parameters as a function of carbon-carbon bond length. 144

6.4 Molecular structures of *anti*-$B_{18}H_{22}$ (a) and *syn*-$B_{18}H_{22}$ (b) isomers. 156

6.5 UV-vis absorption spectra of *anti*-$B_{18}H_{22}$ (a) and *syn*-$B_{18}H_{22}$ (b) and fluorescence emission spectrum (c) (λ_{exc} = 340 nm) of *anti*-$B_{18}H_{22}$ in hexane. 159

6.6 Transient absorption spectra of *anti*-$B_{18}H_{22}$ (a) and *syn*-$B_{18}H_{22}$ (b) excited at 330 nm and recorded at 100 ps and 12 ps, respectively. 161

6.7 Kinetics measured at maxima of transient absorption spectra of *anti*-$B_{18}H_{22}$ (a) and *syn*-$B_{18}H_{22}$ (b) at 512 nm and 475 nm, respectively. 162

6.8 Time dependence of the $O_2(^1\Delta_g)$ phosphorescence at 1270 nm produced by *anti*-$B_{18}H_{22}$ ($\lambda_{exc} = 308$ nm) calculated as the difference of the signal in air- and argon-saturated hexane. Inset: Phosphorescence spectrum of $O_2(^1\Delta_g)$ in an oxygen-saturated chloroform solution of *anti*-$B_{18}H_{22}$ ($\lambda_{exc} = 350$ nm). 163

6.9 Time dependence of the $O_2(^1\Delta_g)$ luminiscence signal at 1275 nm produced by *anti*-$B_{18}H_{22}$ in chloroform ($\lambda_{exc} = 355$ nm): air-saturated (a) and N_2-saturated solution (b). 164

6.10 Most relevant CASSCF molecular orbitals for the *anti*-$B_{18}H_{22}$ isomer. 167

6.11 Most relevant CASSCF molecular orbitals for the *syn*-$B_{18}H_{22}$ isomer. 168

6.12 Photophysics, potential energy hypersurfaces, and structures of the ground state (S_0) and lowest excited singlet state (S_1) for *anti*-$B_{18}H_{22}$ (a) and *syn*-$B_{18}H_{22}$ (b) based on experimental and CASPT2 results. 172

6.13 Atom labels for the systems: (a) *anti*-$B_{18}H_{22}$ molecular geometry at S_0 ground state, C_i symmetry. (b) *syn*-$B_{18}H_{22}$ molecular geometry at S_0 ground state, C_2 symmetry. (c) *anti*-$B_{18}H_{22}$ molecular geometry at $(S_0/S_1)_{CI}$. (d) *syn*-$B_{18}H_{22}$ molecular geometry at $(S_0/S_1)_{CI}$. Significantly elongated and contracted B-B connectivities are shown. 176

6.14 Atom labeling and numbering for the Adenine-Thymine and Guanine-Cytosine base pairs. 181

6.15 Perspective view of the GC base pair in a hydrated DNA double helix of 18 alternating base pairs, namely $d(GC)_9 \cdot d(GC)_9$. . . . 183

6.16 Scheme of the paths explored for the excited state double proton/hydrogen transfer process in the AT base pair. The atoms transferred are H_{61} of the amino group in A plus H'_3 in T. . . . 185

6.17 Scheme of the paths explored for the excited state double proton/hydrogen transfer process in the GC base pair. The atoms transferred are H_1 in G plus H'_{41} of the amino group in C. . . . 186

6.18 Scheme of the paths explored for the excited state double proton/hydrogen transfer process in the GC base pair. The atoms transferred are H_{21} and H'_{41} of the amino groups in both G and C, respectively. 187

6.19 Scheme of the paths explored for the excited state double proton/hydrogen transfer process in the GC base pair. The atoms transferred are H'_{41} in the amino group of C plus either H_1 or H_{21} in G. 188

6.20 SDPT mechanism (H_{61}, H'_3). AT-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1, S_2, S_3, and S_4) are mapped at the CASPT2 level along the O'_4H_{61} and H'_3N_1 reaction coordinates between the WC AT and $TAU_{(AT)}$ structures via the $INT1ION_{(AT)}$ intermediate. 193

6.21 SDPT mechanism (H_1, H'_{41}). GC-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1, S_2, and S_3) are mapped at the CASPT2 level along the N_1H_1 and $N'_4H'_{41}$ reaction coordinates between the WC GC and $TAU1_{(GC)}$ structures via the $INT1ION_{(GC)}$ intermediate. . 194

6.22 SDPT mechanism (H'_3, H_{61}). AT-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1 and S_2) are mapped at the CASPT2 level along the O'_4H_{61} and H'_3N_1 reaction coordinates between the WC AT and $TAU1_{(AT)}$ structures via the $INT2ION_{(AT)}$ intermediate. 195

6.23 SDPT mechanism (H'_{41}, H_1). GC-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1, S_2, and S_3) are mapped at the CASPT2 level along the $N'_4H'_{41}$ and N_1H_1 reaction coordinates between the WC GC and $TAU1_{(GC)}$ structures via the $INT3ION_{(GC)}$ intermediate. . 196

6.24 SDPT mechanism (H'_{41}, H_{21}). GC-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1, S_2, and S_3) are mapped at the CASPT2 level along the $N'_4H'_{41}$ and N_2H_{21} reaction coordinates between the WC GC and $TAU2_{(GC)}$ structures via the $INT3ION_{(GC)}$ intermediate. . 197

6.25 SDHT mechanism (H_{61}, H'_3). AT-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1, S_2, and S_3) are mapped at the CASPT2 level along the O'_4H_{61} and H'_3N_1 reaction coordinates between the WC AT and $TAU_{(AT)}$ structures via the $INT1NEU_{(AT)}$ intermediate. . . 199

6.26 SDHT mechanism (H_1, H'_{41}). GC-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1, S_2, and S_3) are mapped at the CASPT2 level along the N_1H_1 and $N'_4H'_{41}$ reaction coordinates between the WC GC and $TAU1_{(GC)}$ structures via the $INT1NEU_{(GC)}$ intermediate. . 200

6.27 SDHT mechanism (H_{21}, H'_{41}). GC-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1, S_2, and S_3) are mapped at the CASPT2 level along the N_2H_{21} and $N'_4H'_{41}$ reaction coordinates between the WC GC and $TAU2_{(GC)}$ structures via the $INT2NEU_{(GC)}$ intermediate. . 201

6.28 CDPT mechanism (H_{61}+H'_3). AT-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1 and S_2) are mapped at the CASPT2 level along the O'_4H_{61} and H'_3N_1 reaction coordinates between the WC AT and $TAU_{(AT)}$ structures via the concerted pathway. 202

6.29 CDPT mechanism (H_1+H'_{41}). GC-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1, S_2, and S_3) are mapped at the CASPT2 level along the N_1H_1 and $N'_4H'_{41}$ reaction coordinates between the WC GC and $TAU1_{(GC)}$ structures via the concerted pathway. 203

6.30 CDPT mechanism (H_{21}+H'_{41}). GC-base-pair adiabatic PESs of the ground state (S_0) and the low-lying $\pi\pi^*$ singlet excited states (S_1, S_2, and S_3) mapped at the CASPT2 level along the N_2H_{21} and $N'_4H'_{41}$ reaction coordinates between the WC GC and $TAU2_{(GC)}$ structures via the concerted pathway. 204

6.31 SDPT and SDHT mechanisms in the DNA-embedded GC base pair, $d(GC)_9{\cdot}d(GC)_9$, between the canonical WC form and the tautomeric TAU1 and TAU2 species. 206

6.32 Plot of CASPT2 energies, NH distances, and the magnitude of the dipole moments of the S_0 and S_1 states vs. time for a set of semiclassical dynamics simulations. 210

Bibliography

[1] B.O. Roos. *Advances in Chemical Physics, Ab Initio Methods in Quantum Chemistry*, chapter The Complete Active Space Self-Consistent Field Method and its Applications in Electronic Structure Calculations, pages 399–445. John Wiley & Sons, Inc., 1987. 1.1, 3.2, 6.1, 6.1.3

[2] J. Olsen, B.O. Roos, P. Jørgensen, and H.J.Aa. Jensen. Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces. *Journal of Chemical Physics*, 89(4):2185–2192, 1988. 1.1, 6.1

[3] P.-Å. Malmqvist, A. Rendell, and B.O. Roos. The restricted active space self-consistent-field method, implemented with a split graph unitary group approach. *The Journal of Physical Chemistry*, 94(14):5477–5482, 1990. 1.1, 6.1

[4] T. Fleig, J. Olsen, and C.M. Marian. The generalized active space concept for the relativistic treatment of electron correlation. I. Kramers-restricted two-component configuration interaction. *Journal of Chemical Physics*, 114(11):4775–4790, 2001. 1.1

[5] S. Vancoillie. *Ab Initio Calculation and Interpretation of g Factors of transition metal compounds. PhD Thesis.* University of Leuven, Belgium, 2009. 1.1

[6] P.-Å. Malmqvist, K. Pierloot, A.R.M. Shahi, C.J. Cramer, and L. Gagliardi. The restricted active space followed by second-order perturbation theory method: Theory and application to the study of CuO_2 and Cu_2O_2 systems. *Journal of Chemical Physics*, 128(20):204109, 2008. 1.1, 6.1, 6.1.3, 6.2.1

[7] F. Aquilante, L. De Vico, N. Ferré, G. Ghigo, P.-Å. Malmqvist, P. Neogrády, T. B. Pedersen, M. Pitonák, M. Reiher, B.O. Roos, L. Serrano-Andrés, M. Urban, V. Veryazov, and R. Lindh. MOLCAS 7: The Next Generation. *Journal of Computational Chemistry*, 31(1):224–247, 2010. 1.1, 5, 6.1

[8] P. Celani and H.-J. Werner. Multireference perturbation theory for large restricted and selected active space reference wave functions. *Journal of Chemical Physics*, 112(13):5546–5557, 2000. 1.1

[9] K. Andersson, P.-Å. Malmqvist, and B.O. Roos. Second-order perturbation theory with a complete active space self-consistent field reference function. *Journal of Chemical Physics*, 96(2):1218–1226, 1992. 1.1, 3.3, 5, 6.1, 6.3.1, 6.4

[10] B.O. Roos, K. Andersson, M.P. Fülscher, P.-Å. Malmqvist, L. Serrano-Andrés, K. Pierloot, and M. Merchán. *Multiconfigurational Perturbation Theory: Applications in Electronic Spectroscopy.* John Wiley & Sons, Inc., 2007. 1.1, 6.1, 6.1.1, 6.3.1, 6.4

[11] B.O. Roos, K. Andersson, M.P. Fülscher, L. Serrano-Andrés, K. Pierloot, M. Merchán, and V. Molina. Applications of level shift corrected perturbation theory in electronic spectroscopy. *Journal of Molecular Structure: THEOCHEM*, 388(0):257–276, 1996. 1.1, 6.1.3

[12] N. Forsberg and P.-Å. Malmqvist. Multiconfiguration perturbation theory with imaginary level shift. *Chemical Physics Letters*, 274(1-3):196–204, 1997. 1.1, 5, 6.1

[13] K. Andersson and B.O. Roos. Multiconfigurational second-order perturbation theory: A test of geometries and binding energies. *International Journal of Quantum Chemistry*, 45(6):591–607, 1993. 1.1, 6.1

[14] G. Ghigo, B.O. Roos, and P.-Å. Malmqvist. A modified definition of the zeroth-order Hamiltonian in multiconfigurational perturbation theory (CASPT2). *Chemical Physics Letters*, 396(1-3):142–149, 2004. 1.1, 5, 6.1, 6.2.1, 6.2.3

[15] J. Finley, P.-Å. Malmqvist, B.O. Roos, and L. Serrano-Andrés. The multi-state CASPT2 method. *Chemical Physics Letters*, 288(2-4):299–306, 1998. 1.1, 1.1, 6.1, 6.1.2, 6.3, 6.1.2, 6.2.1

[16] L. Serrano-Andrés, M. Merchán, and R. Lindh. Computation of conical intersections by using perturbation techniques. *Journal of Chemical Physics*, 122(10):104107, 2005. 1.1, 5, 5, 6.1.2, 6.1.3, 6.3.1

[17] D. Bohm and D. Pines. A Collective Description of Electron Interactions: III. Coulomb Interactions in a Degenerate Electron Gas. *Physical Review*, 92(3):609–625, 1953. 1.2

[18] M. Gell-Mann and K.A. Brueckner. Correlation Energy of an Electron Gas at High Density. *Physical Review*, 106(2):364–368, 1957. 1.2

[19] J. Hubbard. The description of collective motions in terms of many-body perturbation theory. II. The correlation energy of a free-electron gas. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 243(1234):336–352, 1958. 1.2

[20] J. Hubbard. Electron correlations in narrow energy bands. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 276(1365):238–257, 1963. 1.2, 6.2

[21] J. Quintanilla and C. Hooley. The strong-correlations puzzle. *Physics World Archive*, pages 32–37, 2009. 1.2

[22] A. Altland and B. Simons. *Concepts of theoretical solid state physics*. 2001. 1.2

[23] V.M. García, R. Caballol, and J.P. Malrieu. Theoretical study of the ethylene electronic spectrum and extraction of an r-dependent Hubbard Hamiltonian. *Chemical Physics Letters*, 261(1-2):98–104, 1996. 1.2, 6.2.2

[24] T.G. Schmalz, L. Serrano-Andrés, V. Sauri, M. Merchán, and J.M. Oliva. A distance-dependent parameterization of the extended Hubbard model for conjugated and aromatic hydrocarbons derived from stretched ethene. *The Journal of Chemical Physics*, 135(19):194103, 2011. 1.2, 6.1

[25] G.N. Lewis. The atom and the molecule. *Journal of the American Chemical Society*, 38(4):762–785, 1916. 1.3

[26] E. Schrödinger. Quantisation as an eigen value problem. *Annalen der Physik*, 79(4):361–368, 1926. 1.3

[27] W. Heitler and F. London. Interaction of neutral atoms and covalent bonds in quantum mechanics. *Zeitschrift für Physik*, 44:455–472, 1927. 1.3

[28] L. Pauling. The nature of the chemical bond. Application of results obtained from the quantum mechanics and from a theory of paramagnetic susceptibility to the structure of molecules. *Journal of the American Chemical Society*, 53(4):1367–1400, 1931. 1.3

[29] R.J. Gillespie and R.S. Nyholm. Inorganic stereochemistry. *Quarterly Reviews, Chemical Society*, 11(4):339–380, 1957. 1.3

[30] B.E. Douglas, D.H. McDaniel, and J.J. Alexander. *Conceptos y modelos de Química inorgánica*. Editorial Reverté, 1987. 1.3, 1.3

[31] F.A. Cotton and G. Wilkinson. *Advanced Inorganic Chemistry*. John Wiley & Sons, 1988. 1.3, 1.3

[32] F.A. Kiani and M. Hofmann. Which *nido* : *nido*-macropolyhedral boranes are most stable? *Inorganic Chemistry*, 45(17):6996–7003, 2006. 1.3, 1.3, 1.3

[33] R.W. Rudolph. Boranes and heteroboranes: a paradigm for the electron requirements of clusters? *Accounts of Chemical Research*, 9(12):446–452, 1976. 1.3

[34] K. Wade. Structural and bonding patterns in cluster chemistry. volume 18 of *Advances in Inorganic Chemistry*, pages 1–66. Academic Press, 1976. 1.3

[35] E.D. Jemmis, M.M. Balakrishnarajan, and P.D. Pancharatna. A unifying electron-counting rule for macropolyhedral boranes, metallaboranes, and metallocenes. *Journal of the American Chemical Society*, 123(18):4313–4323, 2001. 1.3

[36] T.D. Getman, J.A. Krause, and S.G. Shore. Synthesis of the new boron hydride *nido*-undecaborane(15), $B_{11}H_{15}$, and the X-ray structure of its conjugate base tetradecahydroundecaborate(1-), $[B_{11}H_{14}]^-$. *Inorganic Chemistry*, 27(14):2398–2399, 1988. 1.3

[37] C.T. Brewer and R.N. Grimes. Metal-induced oxidative fusion of boranes. Synthesis of dodecaborane(16), the first neutral dodecaborane. *Journal of the American Chemical Society*, 106(9):2722–2723, 1984. 1.3

[38] C.T. Brewer, R.G. Swisher, E. Sinn, and R.N. Grimes. Metal-promoted fusion of $B_6H_9^-$. Directed synthesis and structural characterization of dodecaborane(16), $B_{12}H_{16}$. *Journal of the American Chemical Society*, 107(12):3558–3564, 1985. 1.3

[39] A.R. Pitochelli and M.F. Hawthorne. The preparation of a new boron hydride $B_{18}H_{22}$. *Journal of the American Chemical Society*, 84(16):3218–3218, 1962. 1.3, 6.3, 6.3.1

[40] P.G. Simpson and W.N. Lipscomb. Molecular structure of $B_{18}H_{22}$. *Proceedings of the National Academy of Sciences*, 48(9):1490–1491, 1962. 1.3, 6.3.3

[41] P.G. Simpson, K. Folting, R.D. Dobrott, and W. N. Lipscomb. Molecular, crystal, and valence structures of *iso*-$B_{18}H_{22}$. *The Journal of Chemical Physics*, 39(9):2339–2348, 1963. 1.3

[42] P.G. Simpson, K. Folting, and W.N. Lipscomb. The molecular structure of *i*-$B_{18}H_{22}$. *Journal of the American Chemical Society*, 85(12):1879–1880, 1963. 1.3, 1.8, 6.3.3

[43] P.G. Simpson and W.N. Lipscomb. Molecular, crystal, and valence structures of $B_{18}H_{22}$. *The Journal of Chemical Physics*, 39(1):26–34, 1963. 1.3

[44] F.P. Olsen, R.C. Vasavada, and M.F. Hawthorne. The chemistry of *n*-$B_{18}H_{22}$ and *i*-$B_{18}H_{22}$. *Journal of the American Chemical Society*, 90(15):3946–3951, 1968. 1.3

[45] Y. Kawasaki, T. Kuroi, T. Yamashita, K. Horita, T. Hayashi, M. Ishibashi, M. Togawa, Y. Ohno, M. Yoneda, T. Horsky, D. Jacobson, and W. Krull. Ultra-shallow junction formation by $B_{18}H_{22}$ ion implantation. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 237(1-2):25–29, 2005. 1.3, 6.3

[46] K.S. Cook. Methods of preparing clusterboron, 2011. US 20110195009 A1. 1.3

[47] J.D. Watson and F.H.C Crick. Molecular structure of nucleic acids: A structure for Deoxyribose Nucleic Acid. *Nature*, 171(4356):737–738, 1953. 1.4, 1.4, 6.4

[48] J.D. Watson and F.H.C. Crick. Genetical implications of the structure of Deoxyribonucleic Acid. *Nature*, 171(4356):964–967, 1953. 1.4, 1.4

[49] J.D. Watson and F.H.C. Crick. The complementary structure of Deoxyribonucleic Acid. *Proceedings of the Royal Society London A*, 223(1152):80–96, 1954. 1.4, 1.4

[50] P. Schuster and P. Wolschann. Hydrogen bonding: From small clusters to biopolymers. *Monatshefte für Chemie / Chemical Monthly*, 130(8):947–960, 1999. 1.4

[51] R.R. Sinden. *DNA Structure and Function.* Academic Press, San Diego, 1994. 1.4

[52] A. Kornberg and T.A. Baker. *DNA Replication.* University Science Books, 2005. 1.4

[53] R. Langridge, H.R. Wilson, C.W. Hooper, M.H.F. Wilkins, and L.D. Hamilton. The molecular configuration of deoxyribonucleic acid: I. X-ray diffraction study of a crystalline form of the lithium salt. *Journal of Molecular Biology*, 2(1):19–37, 1960. 1.4

[54] P.-O. Löwdin. Proton tunneling in DNA and its biological implications. *Reviews of Modern Physics*, 35(3):724–732, 1963. 1.4, 6.4, 6.4.1

[55] P.-O. Löwdin. Quantum genetics and the aperiodic solid: Some aspects on the biological problems of heredity, mutations, aging, and tumors in view of the quantum theory of the DNA molecule. volume 2 of *Advances in Quantum Chemistry*, pages 213–360. Academic Press, 1966. 1.4, 6.4

[56] R. Rein and F.E. Harris. Studies of hydrogen-bonded systems. I. The electronic structure and the double well potential of the N-H···N hydrogen bond of the Guanine-Cytosine base pair. *The Journal of Chemical Physics*, 41(11):3393–3401, 1964. 1.4

[57] S. Lunell and G. Sperber. Study of the hydrogen bonding in the Adenine-Thymine, Adenine-Cytosine, and Guanine-Thymine base pairs. *The Journal of Chemical Physics*, 46(6):2119–2124, 1967. 1.4

[58] S. Scheiner and C.W. Kern. Theoretical study of proton transfers between base pairs of DNA. *Chemical Physics Letters*, 57(3):331–333, 1978. 1.4

[59] S. Scheiner and C.W. Kern. Molecular orbital investigation of multiply hydrogen bonded systems. Formic acid dimer and DNA base pairs. *Journal of the American Chemical Society*, 101(15):4081–4085, 1979. 1.4

[60] E. Clementi, J. Mehl, and W. von Niessen. Study of the electronic structure of molecules. XII. Hydrogen bridges in the Guanine-Cytosine pair and in the dimeric form of formic acid. *The Journal of Chemical Physics*, 54(2):508–520, 1971. 1.4

[61] E. Clementi. Computation of large molecules with the Hartree-Fock model. *Proceedings of the National Academy of Sciences of the United States of America*, 69(10):2942–2944, 1972. 1.4

[62] Y.S. Kong, M.S. Jhon, and P.-O. Löwdin. Studies on proton transfers in water clusters and DNA base pairs. *International Journal of Quantum Chemistry*, 32(S14):189–209, 1987. 1.4

[63] V. Hrouda, J. Florián, and P. Hobza. Structure, energetics, and harmonic vibrational spectra of the Adenine-Thymine and Adenine*-Thymine* base pairs: Gradient nonempirical and semiempirical study. *The Journal of Physical Chemistry*, 97(8):1542–1557, 1993. 1.4

[64] J. Florián, V. Hrouda, and P. Hobza. Proton transfer in the Adenine-Thymine base pair. *Journal of the American Chemical Society*, 116(4):1457–1460, 1994. 1.4

[65] J. Florián and J. Leszczynski. Spontaneous DNA mutations induced by proton transfer in the Guanine-Cytosine base pairs: An energetic perspective. *Journal of the American Chemical Society*, 118(12):3010–3017, 1996. 1.4, 1.4

[66] A. Douhal, S.K. Kim, and A.H. Zewail. Femtosecond molecular dynamics of tautomerization in model base pairs. *Nature*, 378(6554):260–263, 1995. 1.4, 6.4

[67] M. Chachisvilis, T. Fiebig, A. Douhal, and A.H. Zewail. Femtosecond dynamics of a hydrogen-bonded model base pair in the condensed phase: Double proton transfer in 7-azaindole. *Journal of Physical Chemistry A*, 102(4):669–673, 1998. 1.4, 6.4

[68] A. Nakajima, M. Hirano, R. Hasumi, K. Kaya, H. Watanabe, C.C. Carter, J.M. Williamson, and T.A. Miller. High-resolution laser-induced fluorescence spectra of 7-azaindole-water complexes and 7-azaindole dimer. *The Journal of Physical Chemistry A*, 101(4):392–398, 1997. 1.4

[69] R. López-Martens, P. Long, D. Solgadi, B. Soep, J. Syage, and P. Millie. A time-resolved photoelectron study of the double excited-state proton-transfer reaction in 7-azaindole dimer. *Chemical Physics Letters*, 273(3-4):219–226, 1997. 1.4, 6.4

[70] S. Takeuchi and T. Tahara. Observation of dimer excited-state dynamics in the double proton transfer reaction of 7-azaindole by femtosecond fluorescence up-conversion. *Chemical Physics Letters*, 277(4):340–346, 1997. 1.4, 6.4

[71] S. Takeuchi and T. Tahara. Femtosecond ultraviolet-visible fluorescence study of the excited-state proton-transfer reaction of 7-azaindole dimer. *Journal of Physical Chemistry A*, 102(40):7740–7753, 1998. 1.4

[72] D.E. Folmer, L. Poth, E.S. Wisniewski, and A.W. Castleman Jr. Arresting intermediate states in a chemical reaction on a femtosecond time scale: Proton transfer in model base pairs. *Chemical Physics Letters*, 287(1-2):1–7, 1998. 1.4, 6.4

[73] D.E. Folmer, E.S. Wisniewski, and A.W. Castleman Jr. Excited state double proton transfer in the 7-azaindole dimer revisited. *Chemical Physics Letters*, 318(6):637–643, 2000. 1.4

[74] O.-H. Kwon and A.H. Zewail. Double proton transfer dynamics of model DNA base pairs in the condensed phase. *Proceedings of the National Academy of Sciences*, 104(21):8703–8708, 2007. 1.4, 6.4

[75] A. Douhal, V. Guallar, M. Moreno, and J.M. Lluch. Theoretical study of molecular dynamics in model base pairs. *Chemical Physics Letters*, 256(4-5):370–376, 1996. 1.4, 6.4

[76] M. Moreno, A. Douhal, J.M. Lluch, O. Castaño, and L.M. Frutos. *Ab initio* based exploration of the potential energy surface for the double proton transfer in the first excited singlet electronic state of the 7-azaindole dimer. *The Journal of Physical Chemistry A*, 105(15):3887–3893, 2001. 1.4, 6.4

[77] J. Catalán, J.C. del Valle, and M. Kasha. Resolution of concerted versus sequential mechanisms in photo-induced double-proton transfer reaction in 7-azaindole H-bonded dimer. *Proceedings of the National Academy of Sciences of the United States of America*, 96(15):8338–8343, 1999. 1.4, 6.4

[78] V. Guallar, A. Douhal, M. Moreno, and J.M. Lluch. DNA mutations induced by proton and charge transfer in the low-lying excited singlet electronic states of the DNA base pairs: A theoretical insight. *The Journal of Physical Chemistry A*, 103(31):6251–6256, 1999. 1.4

[79] L. Serrano-Andrés and M. Merchán. Theoretical CASPT2 study of the excited state double proton transfer reaction in the 7-azaindole dimer. *Chemical Physics Letters*, 418(4-6):569–575, 2006. 1.4, 6.4

[80] A.L. Sobolewski and W. Domcke. *Ab initio* studies on the photophysics of the Guanine-Cytosine base pair. *Physical Chemistry Chemical Physics*, 6(10):2763–2771, 2004. 1.4, 6.4

[81] A.L. Sobolewski, W. Domcke, and C. Hättig. Tautomeric selectivity of the excited-state lifetime of Guanine/Cytosine base pairs: The role of electron-driven proton-transfer processes. *Proceedings of the National Academy of Sciences of the United States of America*, 102(50):17903–17906, 2005. 1.4, 6.4, 6.4.1, 6.4.1

[82] S. Perun, A.L. Sobolewski, and W. Domcke. Role of electron-driven proton-transfer processes in the excited-state deactivation of the Adenine-Thymine base pair. *The Journal of Physical Chemistry A*, 110(29):9031–9038, 2006. PMID: 16854013. 1.4, 6.4, 6.4.1

[83] L.M. Frutos, A. Markmann, A.L. Sobolewski, and W. Domcke. Photoinduced electron and proton transfer in the hydrogen-bonded pyridine-pyrrole system. *The Journal of Physical Chemistry B*, 111(22):6110–6112, 2007. 1.4, 6.4, 6.4.1

[84] A. Abo-Riziq, L. Grace, E. Nir, M. Kabelac, P. Hobza, and M.S. de Vries. Photochemical selectivity in Guanine-Cytosine base-pair structures. *Proceedings of the National Academy of Sciences of the United States of America*, 102(1):20–23, 2005. 1.4, 6.4.1

[85] N.K. Schwalb and F. Temps. Ultrafast electronic relaxation in Guanosine is promoted by hydrogen bonding with Cytidine. *Journal of the American Chemical Society*, 129(30):9272–9273, 2007. 1.4, 6.4

[86] F.-A. Miannay, A. Bányász, T. Gustavsson, and D. Markovitsi. Ultrafast excited-state deactivation and energy transfer in Guanine-Cytosine DNA double helices. *Journal of the American Chemical Society*, 129(47):14574–14575, 2007. 1.4, 6.4, 6.4.2, 6.4.3, 7

[87] K. de La Harpe, C.E. Crespo-Hernández, and B. Kohler. Deuterium isotope effect on excited-state dynamics in an alternating GC oligonucleotide. *Journal of the American Chemical Society*, 131(48):17557–17559, 2009. 1.4, 7

[88] C.E. Crespo-Hernández, K. de La Harpe, and B. Kohler. Ground-state recovery following UV excitation is much slower in G·C-DNA duplexes and hairpins than in mononucleotides. *Journal of the American Chemical Society*, 130(33):10844–10845, 2008. 1.4, 6.4, 6.4.2, 6.4.3

[89] G. Groenhof, L.V. Schäfer, M. Boggio-Pasqua, M. Goette, H. Grubmüller, and M.A. Robb. Ultrafast deactivation of an excited Cytosine-Guanine base pair in DNA. *Journal of the American Chemical Society*, 129(21):6812–6819, 2007. 1.4, 6.4, 6.4.3, 6.4.3, 6.4.3, 7

[90] I. Vayá, F.-A. Miannay, T. Gustavsson, and D. Markovitsi. High-energy long-lived excited states in DNA double strands. *ChemPhysChem*, 11(5):987–989, 2010. 1.4, 6.4

[91] I. Vayá, P. Changenet-Barret, T. Gustavsson, D. Zikich, A.B. Kotlyar, and D. Markovitsi. Long-lived fluorescence of homopolymeric Guanine-Cytosine DNA duplexes. *Photochemical and Photobiological Sciences*, 9(9):1193–1195, 2010. 1.4, 6.4

[92] L. Biemann, S.A. Kovalenko, K. Kleinermanns, R. Mahrwald, M. Markert, and R. Improta. Excited state proton transfer is not involved in the ultrafast deactivation of Guanine-Cytosine pair in solution. *Journal of the American Chemical Society*, 133(49):19664–19667, 2011. 1.4, 6.4.2, 7

[93] E.S. Kryachko. *Fundamental World of Quantum Chemistry*. Kluwer Academic Publishers, 2003. 1.11, 1.12, 1.13

[94] E.S. Kryachko and J.R. Sabin. Quantum chemical study of the hydrogen-bonded patterns in A·T base pair of DNA: Origins of tautomeric mispairs, base flipping, and Watson-Crick: Hoogsteen conversion. *International Journal of Quantum Chemistry*, 91(6):695–710, 2003. 1.4

[95] A. Szabo and N.S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, 1996. 3, 3, 3, 3.1, 3.1, 3.1, 3.3, 3.3, 3.4, 4, 5, 5

[96] M. Merchán and L. Serrano-Andrés. *Computational Photochemistry*, volume 16, chapter 2, page 35. Elsevier, 2005. 3, 3.3, 5, 5

[97] T. Helgaker, P. Jørgensen, and J. Olsen. *Molecular Electronic-Structure Theory*. John Wiley & Sons, Inc., 2004. 3, 3.2, 3.3, 5

[98] P.-O. Widmark and B.O. Roos. *European Summerschool of Quantum Chemistry 2005*, volume Books I, II, III. Lund University, 2005. 3, 3.2, 5

[99] F. Jensen. *Introduction to Computational Chemistry*. John Wiley & Sons, Inc., 1999. 3, 3.4, 4, 6.2.3

[100] J.A. Gaunt. A theory of Hartree's atomic fields. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24(02):328–342, 1928. 3.1

[101] V.A. Fock. *Zeitschrift für Physik*, 15:126, 1930. 3.1

[102] C.C.J. Roothaan. New developments in Molecular Orbital theory. *Reviews of Modern Physics*, 23(2):69–89, 1951. 3.1, 4

[103] A. Meckler. Electronic energy levels of molecular Oxygen. *The Journal of Chemical Physics*, 21(10):1750–1762, 1953. 3.2

[104] B.O. Roos, P.R. Taylor, and P.E.M. Siegbahn. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chemical Physics*, 48(2):157, 1980. 3.2

[105] B.O. Roos. The complete active space SCF method in a Fock-matrix-based super-CI formulation. *International Journal of Quantum Chemistry, Quantum Chemistry Symposium*, 14:175–189, 1980. 3.2

[106] B. Huron, J.P. Malrieu, and P. Rancurel. Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wave functions. *The Journal of Chemical Physics*, 58(12):5745–5759, 1973. 3.2

[107] R. Buenker and S. Peyerimhoff. Individualized configuration selection in CI calculations with subsequent energy extrapolation. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, 35(1):33–58, 1974. 3.2

[108] C. Møller and M.S. Plesset. Note on an approximation treatment for many-electron systems. *Physical Review*, 46(7):618–622, 1934. 3.3, 6.2.1, 6.2.3

[109] K. Andersson, P.-Å. Malmqvist, B.O. Roos, A.J. Sadlej, and K. Wolinski. Second-order perturbation theory with a CASSCF reference function. *The Journal of Physical Chemistry*, 94(14):5483–5488, 1990. 3.3, 5, 6.3.1

[110] J. Čížek. On the correlation problem in atomic and molecular systems. calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods. *The Journal of Chemical Physics*, 45(11):4256–4266, 1966. 3.4

[111] J. Čížek and J. Paldus. Correlation problems in atomic and molecular systems III. Rederivation of the coupled-pair many-electron theory using the traditional quantum chemical methods. *International Journal of Quantum Chemistry*, 5(4):359–379, 1971. 3.4

[112] J. Čížek. *On the Use of the Cluster Expansion and the Technique of Diagrams in Calculations of Correlation Effects in Atoms and Molecules*, pages 35–89. John Wiley & Sons, Inc., 2007. 3.4

[113] R.J. Bartlett. Coupled-cluster approach to molecular structure and spectra: a step toward predictive quantum chemistry. *The Journal of Physical Chemistry*, 93(5):1697–1708, 1989. 3.4

[114] P. Sherwood. *Modern Methods and Algorithms of Quantum Chemistry, Proceedings*, volume 3, chapter 285-305. John von Newmann Institute for Computing, Jülich, 2000. 3.5, 3.3

[115] H.M. Senn and W. Thiel. QM/MM methods for biomolecular systems. *Angewandte Chemie International Edition*, 48(7):1198–1229, 2009. 3.5, 3.5

[116] A.C.T. van Duin, S. Dasgupta, F. Lorant, and W.A. Goddard. ReaxFF: A reactive force field for hydrocarbons. *The Journal of Physical Chemistry A*, 105(41):9396–9409, 2001. 3.5

[117] H. Lin and D. Truhlar. QM/MM: what have we learned, where are we, and where do we go from here? *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, 117(2):185–199, 2007. 3.5

[118] G. Stock and M. Thoss. *Classical Description of Nonadiabatic Quantum Dynamics*, pages 243–375. John Wiley & Sons, Inc., 2005. 3.6

[119] J.C. Tully and R.K. Preston. Trajectory surface hopping approach to nonadiabatic molecular collisions: The reaction of H^+ with D_2. *The Journal of Chemical Physics*, 55(2):562–572, 1971. 3.6

[120] J.C. Tully. Molecular dynamics with electronic transitions. *The Journal of Chemical Physics*, 93(2):1061–1071, 1990. 3.6

[121] U. Müller and G. Stock. Surface-hopping modeling of photoinduced relaxation dynamics on coupled potential-energy surfaces. *The Journal of Chemical Physics*, 107(16):6230–6245, 1997. 3.6

[122] T. Vreven, F. Bernardi, M. Garavelli, M. Olivucci, M.A. Robb, and H.B. Schlegel. *Ab initio* photoisomerization dynamics of a simple retinal chromophore model. *Journal of the American Chemical Society*, 119(51):12687–12688, 1997. 3.6

[123] G. Granucci, M. Persico, and A. Toniolo. Direct semiclassical simulation of photochemical processes with semiempirical wave functions. *The Journal of Chemical Physics*, 114(24):10608–10615, 2001. 3.6

[124] M. Barbatti, G. Granucci, M. Persico, M. Ruckenbauer, M. Vazdar, M. Eckert-Maksić, and H. Lischka. The on-the-fly surface-hopping program system Newton-X: Application to *ab initio* simulation of the nonadiabatic photodynamics of benchmark systems. *Journal of Photochemistry and Photobiology A: Chemistry*, 190(2-3):228–240, 2007. 3.6, 5

[125] L. Verlet. Computer experiments on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Physical Review*, 159(1):98–103, 1967. 3.6, 6.4

[126] J.C. Slater. Atomic shielding constants. *Physical Review*, 36(1):57–64, 1930. 4

[127] S.F. Boys. Electronic wave functions. I. A general method of calculation for the stationary states of any molecular system. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 200(1063):542–554, 1950. 4

[128] J. Almlöf and P.R. Taylor. General contraction of Gaussian basis sets. I. Atomic natural orbitals for first- and second-row atoms. *The Journal of Chemical Physics*, 86(7):4070–4077, 1987. 4, 4

[129] W.J. Hehre, R. Ditchfield, and J.A. Pople. Self-consistent molecular orbital methods. XII. Further extensions of Gaussian-type basis sets for use in molecular orbital studies of organic molecules. *The Journal of Chemical Physics*, 56(5):2257–2261, 1972. 4

[130] R. Krishnan, J.S. Binkley, R. Seeger, and J.A. Pople. Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *The Journal of Chemical Physics*, 72(1):650–654, 1980. 4

[131] S. Huzinaga. Gaussian-type functions for polyatomic systems. I. *The Journal of Chemical Physics*, 42(4):1293–1302, 1965. 4

[132] F.B. van Duijneveldt. *Technical Report, IBM Res. Rep. RJ945*, 445, 1971. 4

[133] H. Partridge. Near Hartree-Fock quality GTO basis sets for the second-row atoms. *The Journal of Chemical Physics*, 87(11):6643–6647, 1987. 4

[134] K. Pierloot, B. Dumez, P.-O. Widmark, and B.O. Roos. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, 90(2):87–114, 1995. 4, 6.4

[135] M. Merchán, L. Serrano-Andrés, M.P. Fülscher and B.O. Roos. *Recent Advances in Multireference Theory*, volume IV, chapter Multiconfigurational Perturbation Theory Applied to Excited States of Organic Compounds, pages 161–195. World Scientific Publishing, 1999. 4

[136] B.O. Roos, M. P. Fülscher, P.-Å. Malmqvist, M. Merchán, and L. Serrano-Andrés. *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*, pages 357–431. Kluwer, Dordrecht, The Netherlands, 1995. 4

[137] B.O. Roos, K. Andersson, M. P. Fülscher, P.-Å. Malmqvist, L. Serrano-Andrés, K. Pierloot, and M. Merchán. *Advances in Chemical Physics: New Methods in Computational Quantum Mechanics*, volume XCIII, pages 219–331. Wiley, New York, 1996. 4

[138] B.O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov, and P.-O. Widmark. Main group atoms and dimers studied with a new relativistic ANO basis set. *The Journal of Physical Chemistry A*, 108(15):2851–2858, 2004. 4

[139] T.H. Dunning Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *The Journal of Chemical Physics*, 90(2):1007–1023, 1989. 4, 6.2.1, 6.2.3

[140] A.K. Wilson, T. van Mourik, and Jr. T.H. Dunning. Gaussian basis sets for use in correlated molecular calculations. VI. Sextuple zeta correlation consistent basis sets for boron through neon. *Journal of Molecular Structure: THEOCHEM*, 388(0):339–349, 1996. 4

[141] R.A. Kendall, Jr. T.H. Dunning, and R.J. Harrison. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *The Journal of Chemical Physics*, 96(9):6796–6806, 1992. 4

[142] D.E. Woon and Jr. T.H. Dunning. Gaussian basis sets for use in correlated molecular calculations. V. Core-valence basis sets for boron through neon. *The Journal of Chemical Physics*, 103(11):4572–4585, 1995. 4

[143] A.K. Wilson and Jr. T.H. Dunning. Benchmark calculations with correlated molecular wave functions. X. Comparison with "exact" MP2 calculations on Ne, HF, H_2O, and N_2. *The Journal of Chemical Physics*, 106(21):8718–8726, 1997. 4

[144] D. Feller and K.A. Peterson. An examination of intrinsic errors in electronic structure methods using the Environmental Molecular Sciences Laboratory computational results database and the Gaussian-2 set. *The Journal of Chemical Physics*, 108(1):154–176, 1998. 4

[145] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A.K. Wilson. Basis-set convergence in correlated calculations on Ne, N_2, and H_2O. *Chemical Physics Letters*, 286(3-4):243–252, 1998. 4

[146] D. Feller and K.A. Peterson. Probing the limits of accuracy in electronic structure calculations: Is theory capable of results uniformly better than "chemical accuracy"? *The Journal of Chemical Physics*, 126(11):114105, 2007. 4, 6.2.1, 6.2.3

[147] A. Requena and J. Zú niga. *Espectroscopia*. Pearson Prentice-Hall, 2004. 5

[148] D. Skoog, F. James, and T. Nieman. *Principios de Análisis Instrumental*. McGraw-Hill, 2000. 5

[149] I.N. Levine. *Espectroscopia Molecular*. Editorial AC, 1980. 5, 5

[150] J. Planelles, I. Climente, and J.G.Díaz. *Espectroscòpia*. Publicacions de la Universitat Jaume I, 2002. 5

[151] D.C. Harris and M.D. Bertolucci. *Symmetry and Spectroscopy. An Introduction to Vibrational and Electronic Spectroscopy*. Dover Publications, 1989. 5

[152] E.B. Wilson Jr., J.C. Decius, and P.C. Cross. *Molecular Vibrations. The Theory of Infrared and Raman Vibrational Spectra*. Dover Publications, 1980. 5

[153] L. Serrano-Andrés. *Estudio teórico del espectro electrónico de sistemas orgánicos*. Tesis doctoral. Departamento de Química Física. Universidad de Valencia, 1994. 5

[154] D.C. Harris. *Análisis Químico Cuantitativo*. Grupo Editorial Iberoamérica, 1992. 5

[155] R.S. Mulliken. Band spectra and Chemistry. *Chemical Reviews*, 6(4):503–545, 1930. 5

[156] L. Smart and E. Moore. *Química del estado sólido: Una introducción*. Addison-Wesley Iberoamericana, 1995. 5

[157] D.F. Shriver, P.W. Atkins, and C.H. Langford. *Química Inorgánica*. Vol. 1,2. Reverté, 1998. 5

[158] K. Wittel and S.P. McGlynn. The orbital concept in molecular spectroscopy. *Chemical Reviews*, 77(5):745–771, 1977. 5, 5

[159] J.Andrés and J.Bertrán, editors. *Química Teórica y Computacional*. Colección Ciencias Experimentales. Universitat Jaume I, 2000. 5

[160] A. Dreuw and M. Head-Gordon. Single-reference *ab initio* methods for the calculation of excited states of large molecules. *Chemical Reviews*, 105(11):4009–4037, 2005. 5

[161] L. Serrano-Andrés and M. Merchán. Vida y luz. Una perspectiva químico-cuántica. *Anales de Química*, 100(3):16, 2004. 5

[162] M. Klessinger and J. Michl. *Excited States and Photochemistry of Organic Molecules*. VCH, New York, 1995. 5, 5, 5, 5

[163] J.L. McHale. *Molecular Spectroscopy*. Prentice Hall, 1999. 5

[164] L. Serrano-Andrés and M. Merchán. *Encyclopedia of Computational Chemistry*, chapter Spectroscopy Applications. Wiley, Chichester, 2004. 5, 5

[165] F. Bernardi, M. Olivucci, and M.A. Robb. Following reaction paths in organic photochemistry: The special role of surface crossings. *Pure and Applied Chemistry*, 67(1):17–24, 1995. 5, 5

[166] H. Lischka, M. Dallos, P.G. Szalay, D.R. Yarkony, and R. Shepard. Analytic evaluation of nonadiabatic coupling terms at the MR-CI level. I. Formalism. *The Journal of Chemical Physics*, 120(16):7322–7329, 2004. 5, 5

[167] N. Koga and K. Morokuma. Determination of the lowest energy point on the crossing seam between two potential surfaces using the energy gradient. *Chemical Physics Letters*, 119(5):371–374, 1985. 5, 5

[168] N.J. Turro. *Modern Molecular Photochemistry*. University Science Books, 1991. 5, 5, 5

[169] S.J. Strickler and R.A. Berg. Relationship between absorption intensity and fluorescence lifetime of molecules. *The Journal of Chemical Physics*, 37(4):814–822, 1962. 5, 5, 6.3.3

[170] O. Rubio-Pons, L. Serrano-Andrés, and M. Merchán. A theoretical insight into the photophysics of acridine. *The Journal of Physical Chemistry A*, 105(42):9664–9673, 2001. 5, 5

[171] O. Rubio-Pons. *Charge-transfer excitations and photophysical properties of molecular building blocks.* KTH Royal Institute of Technology, Stockholm, Sweden, 2005. 5

[172] D.L. Dexter. A theory of sensitized luminescence in solids. *The Journal of Chemical Physics*, 21(5):836–850, 1953. 5

[173] T. Föster. Zwischenmolekulare energiewanderung und fluoreszenz. *Annalen der Physik*, 437(1-2):55–75, 1948. 5

[174] L. Serrano-Andrés and J.J Serrano-Pérez. *Handbook of Computational Chemistry*, chapter Calculation of Excited States: Molecular Photophysics and Photochemistry on Display, pages 483–570. Springer Dordrecht Heidelberg London New York, 2012. 5

[175] R. Bonnett. *Chemical Aspects of Photodynamic Therapy.* Gordon & Breach Science, Amsterdam, 2000. 5

[176] R. Bonnett. Photosensitizers of the porphyrin and phtalocianine series for photodynamic therapy. *Chemmical Society Reviews*, 24:19, 1995. 5, 6.3.3

[177] A. Farazdel and M. Dupuis. On the determination of the minimum on the crossing seam of two potential energy surfaces. *Journal of Computational Chemistry*, 12(2):276–282, 1991. 5

[178] D.R. Yarkony. Systematic determination of intersections of potential energy surfaces using a Lagrange multiplier constrained procedure. *The Journal of Physical Chemistry*, 97(17):4407–4412, 1993. 5

[179] M.R. Manaa and D.R. Yarkony. On the intersection of two potential energy surfaces of the same symmetry. Systematic characterization using a Lagrange multiplier constrained procedure. *The Journal of Chemical Physics*, 99(7):5251–5256, 1993. 5

[180] J.M. Anglada and J.M. Bofill. A reduced-restricted-quasi-Newton-Raphson method for locating and optimizing energy crossing points between two potential energy surfaces. *Journal of Computational Chemistry*, 18(8):992–1003, 1997. 5

[181] I.N. Ragazos, M.A. Robb, F. Bernardi, and M. Olivucci. Optimization and characterization of the lowest energy point on a conical intersection

using an MC-SCF Lagrangian. *Chemical Physics Letters*, 197(3):217–223, 1992. 5

[182] M.J. Bearpark, M.A. Robb, and H.B. Schlegel. A direct method for the location of the lowest energy point on a potential surface crossing. *Chemical Physics Letters*, 223(3):269–274, 1994. 5

[183] B.O. Roos, P. Linse, P.E.M. Siegbahn, and M.R.A. Blomberg. A simple method for the evaluation of the second-order-perturbation energy from external double-excitations with a casscf reference wave function. *Chemical Physics*, 66(1-2):197–207, 1982. 5

[184] F. Aquilante, P.-Å. Malmqvist, T. B. Pedersen, A. Ghosh, and B.O. Roos. Cholesky decomposition-based multiconfiguration second-order perturbation theory (CD-CASPT2): Application to the spin-state energetics of CoIII(diiminato)(NPh). *Journal of Chemical Theory and Computation*, 4(5):694–702, 2008. 5, 6.1, 6.2.1, 6.2.3

[185] M. Douglas and N.M. Kroll. Quantum electrodynamical corrections to the fine structure of Helium. *Annals of Physics*, 82(1):89 – 155, 1974. 5, 6.1

[186] B.A. Hess. Relativistic electronic-structure calculations employing a two-component no-pair formalism with external-field projection operators. *Physical Review A*, 33(6):3742–3748, 1986. 5, 6.1

[187] B.A. Heβ, C.M. Marian, U. Wahlgren, and O. Gropen. A mean-field spin-orbit method applicable to correlated wave functions. *Chemical Physics Letters*, 251(5-6):365 – 371, 1996. 5

[188] P.-Å. Malmqvist and B.O. Roos. The CASSCF state interaction method. *Chemical Physics Letters*, 155(2):189 – 194, 1989. 5

[189] J.W. Ponder. *TINKER - Software Tools for Molecular Design, version 4.2*. Department of biochemistry and molecular biophysics, Washington University, Washington, New York, 2004. 5

[190] X.-J. Lu and W.K. Olson. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Research*, 31(17):510–5121, 2003. 5

[191] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery Jr., T. Vreven, K.N. Kudin, J.C. Burant, J.M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J.E. Knox, H.P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, P.Y. Ayala, K. Morokuma, G.A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, A.D. Daniels, M.C. Strain, O. Farkas, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J.V. Ortiz, Q. Cui, A.G. Baboul, S. Clifford, J. Cioslowski, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, C. Gonzalez, and J.A. Pople. Gaussian 03, Revision C.02. Gaussian, Inc., Wallingford, CT, 2004. 5, 6.10

[192] G. Schaftenaar and J.H. Noordik. Molden: a pre- and post-processing program for molecular and electronic structures. *Journal of Computer-Aided Molecular Design*, 14(2):123–134, 2000. 5

[193] A.-R. Allouche. Gabedit-A graphical user interface for computational chemistry softwares. *Journal of Computational Chemistry*, 32(1):174–182, 2011. 5

[194] S.M. Huber, A.R.M. Shahi, F. Aquilante, C.J. Cramer, and L. Gagliardi. What active space adequately describes Oxygen activation by a late transition metal? : CASPT2 and RASPT2 applied to intermediates from the reaction of O-2 with a Cu(I)-alpha-ketocarboxylate. *Journal of Chemical Theory and Computation*, 5(11):2967–2976, 2009. 6.1

[195] A.R.M. Shahi, C.J. Cramer, and L. Gagliardi. Second-order perturbation theory with complete and restricted active space reference functions applied to oligomeric unsaturated hydrocarbons. *Physical Chemistry Chemical Physics*, 11:10964–10972, 2009. 6.1

[196] F. Ruipérez, F. Aquilante, J.M. Ugalde, and I. Infante. Complete vs. Restricted Active Space Perturbation Theory calculation of the Cr_2 potential energy surface. *Journal of Chemical Theory and Computation*, 7(6):1640–1646, 2011. 6.1

[197] S. Vancoillie, H. Zhao, V.T. Tran, M.F.A. Hendrickx, and K. Pierloot. Multiconfigurational Second-Order Perturbation Theory Restricted Active Space (RASPT2). Studies on mononuclear first-row transition-metal systems. *Journal of Chemical Theory and Computation*, 7(12):3961–3977, 2011. 6.1

[198] K. Pierloot, H. Zhao, and S. Vancoillie. Copper corroles: the question of noninnocence. *Inorganic Chemistry*, 49(22):10316–10329, 2010. 6.1

[199] H. Zhao, K. Pierloot, E.H.G. Langner, J.C. Swarts, J. Conradie, and A. Ghosh. Low-energy states of Manganese-oxo corrole and corrolazine: Multiconfiguration reference *ab initio* calculations. *Inorganic Chemistry*, 51(7):4002–4006, 2012. 6.1

[200] J. Su, W.H.E. Schwarz, and J. Li. Electronic spectra and excited states of Neptunyl and its $[NpO_2Cl_4]^{2-}$ complex. *Inorganic Chemistry*, 51(5):3231–3238, 2012. 6.1

[201] D. Escudero and L. González. RASPT2/RASSCF vs. Range-Separated/Hybrid DFT methods: Assessing the excited states of a Ru(II)bipyridyl complex. *Journal of Chemical Theory and Computation*, 8(1):203–213, 2012. 6.1

[202] M. Radoń and E. Broclawik. Mono- and dinitrosyls on Copper(I) site in a Zeolite model: Effects of static correlation. *The Journal of Physical Chemistry A*, 115(42):11761–11774, 2011. 6.1

[203] P.-O. Widmark, P.-Å. Malmqvist, and B.O. Roos. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, 77(5):291–306, 1990. 6.1, 6.2.1

[204] B.O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov, and P.-O. Widmark. New relativistic ANO basis sets for transition metal atoms. *The Journal of Physical Chemistry A*, 109(29):6575–6579, 2005. 6.1

[205] G. Herzberg. *Molecular Spectra and Molecular Structure. Electronic Spectra and Electronic Structure of Polyatomic Molecules*, volume 3. Van Nostrand, Princeton, NJ, 1966. 6.1, 6.2.1

[206] S. Vancoillie and K. Pierloot. Multiconfigurational g tensor calculations as a probe for the covalency of the copper-ligand bonds in Copper(II) complexes: $[CuCl_4]^{2-}$, $[Cu(NH_3)_4]^{2+}$, and Plastocyanin. *The Journal of Physical Chemistry A*, 112(17):4011–4019, 2008. 6.1, 6.1.3, 6.7

[207] M. Merchán and E. Ortí amd B.O. Roos. Theoretical determination of the electronic spectrum of free base porphin. *Chemical Physics Letters*, 226(1-2):27–36, 1994. 6.1, 6.1.1, 6.1

[208] L. Serrano-Andrés, M. Merchán, M. Rubio, and B.O. Roos. Interpretation of the electronic absorption spectrum of free base porphin by using multiconfigurational second-order perturbation theory. *Chemical Physics Letters*, 295(3):195–203, 1998. 6.1, 6.1.1, 6.1

[209] S.R. Gwaltney and R.J. Bartlett. Coupled-cluster calculations of the electronic excitation spectrum of free base porphin in a polarized basis. *The Journal of Chemical Physics*, 108(16):6790–6798, 1998. 6.1

[210] L. Serrano-Andrés, M. Merchán, I. Nebot-Gil, R. Lindh, and B.O. Roos. Towards an accurate molecular orbital theory for excited states: ethene, butadiene, and hexatriene. *The Journal of Chemical Physics*, 98(4):3151–3162, 1993. 6.1.2, 6.3, 6.1.2, 6.4, 6.2.1, 6.2.3, 6.3.1

[211] L. Serrano-Andrés, J. Sánchez-Marín, and I. Nebot-Gil. Theoretical study of the low-lying states of trans-1,3-butadiene. *The Journal of Chemical Physics*, 97(10):7499–7506, 1992. 6.1.2

[212] O. Christiansen, H. Koch, A. Halkier, P. Jørgensen, T. Helgaker, and A. Sánchez de Merás. Large-scale calculations of excitation energies in coupled-cluster theory: The singlet excited states of benzene. *The Journal of Chemical Physics*, 105(16):6921–6939, 1996. 6.1.2, 6.5

[213] K. Andersson and B.O. Roos. Excitation energies in the nickel atom studied with the complete active space SCF method and second-order perturbation theory. *Chemical Physics Letters*, 191(6):507–514, 1992. 6.1.3

[214] K. Pierloot. *Computational Organometallic Chemistry*, pages 123–158. Marcel Dekker, Inc., New York, 2001. 6.1.3

[215] K. Pierloot. The CASPT2 method in inorganic electronic spectroscopy: from ionic transition metal to covalent actinide complexes. *Molecular Physics*, 101(13):2083–2094, 2003. 6.1.3

[216] S. Vancoillie, J. Chalupský, U. Ryde, E.I. Solomon, K. Pierloot, F. Neese, and L. Rulisek. Multireference *ab initio* calculations of g tensors for trinuclear copper clusters in multicopper oxidases. *The Journal of Physical Chemistry B*, 114(22):7692–7702, 2010. 6.1.3

[217] P.J. Linstrom and W.G. Mallard. *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*. National Institute of Standards and Technology: Gaithersburg MD, 20899, 2010. 6.6

[218] S. Vancoillie, P.-Å. Malmqvist, and K. Pierloot. Calculation of EPR g tensors for transition-metal complexes based on multiconfigurational perturbation theory (CASPT2). *ChemPhysChem*, 8(12):1803–1815, 2007. 6.1.3, 6.7

[219] J.C. Phillips. *Physics of High-Tc Superconductors*. Academic, San Diego, 1989. 6.2

[220] W.P. Su. Lattice relaxation of even-parity singlet excited states in polyacetylene and four-soliton bound state. *Physical Review Letters*, 74(7):1167–1170, 1995. 6.2

[221] V.I. Anisimov, J. Zaanen, and O.K. Andersen. Band theory and mott insulators: Hubbard U instead of stoner I. *Physical Review B*, 44(3):943–954, 1991. 6.2

[222] V.I. Anisimov, F. Aryasetiawan, and A.I. Lichtenstein. First-principles calculations of the electronic structure and spectra of strongly correlated systems: the LDA+ U method. *Journal of Physics: Condensed Matter*, 9(4):767–808, 1997. 6.2

[223] C. Cao, S. Hill, and H.-P. Cheng. Strongly correlated electrons in the $[\mathrm{Ni(hmp)}(R\mathrm{OH})\mathrm{X}]_4$ single molecule magnet: A DFT$+U$ study. *Physical Review Letters*, 100:167206, Apr 2008. 6.2

[224] E. Scriven and B.J. Powell. Toward the parametrization of the Hubbard model for salts of bis(ethylenedithio)tetrathiafulvalene: A density functional study of isolated molecules. *The Journal of Chemical Physics*, 130(10):104508, 2009. 6.2

[225] S. Gangopadhyay, A.E. Masunov, E. Poalelungi, and M.N. Leuenberger. Weak antiferromagnetic coupling in molecular ring is predicted correctly by density functional theory plus hubbard U. *The Journal of Chemical Physics*, 132(24):244104, 2010. 6.2

[226] F.A. Matsen. Correlation of molecular orbital and valence bond states in π systems. *Accounts of Chemical Research*, 11(10):387–392, 1978. 6.2

[227] J. s. Ryu and B.S. Hudson. A new interpretation of the electronic spectrum of ethylene from 6-8 eV. *Chemical Physics Letters*, 245(4-5):448–454, 1995. 6.2.1

[228] A.J. Merer and R.S. Mulliken. Ultraviolet spectra and excited states of ethylene and its alkyl derivatives. *Chemical Reviews*, 69(5):639–656, 1969. 6.2.1

[229] C. Petrongolo, R.J. Buenker, and S.D. Peyerimhoff. Nonadiabatic treatment of the intensity distribution in the V-N bands of ethylene. *The Journal of Chemical Physics*, 76(7):3655–3667, 1982. 6.2.1, 6.2.1, 6.11, 6.2.3

[230] L.E. McMurchie and E.R. Davidson. Singlet Rydberg states of ethylene. *The Journal of Chemical Physics*, 67(12):5613–5618, 1977. 6.2.1, 6.2.1, 6.11, 6.2.3

[231] R. Lindh and B.O. Roos. A theoretical study of the diffuseness of the V($^1B_{1u}$) state of planar ethylene. *International Journal of Quantum Chemistry*, 35(6):813–825, 1989. 6.2.1, 6.2.1, 6.11, 6.2.3

[232] V. Sauri, L. Serrano-Andrés, A.R. M. Shahi, L. Gagliardi, S. Vancoillie, and K. Pierloot. Multiconfigurational second-order perturbation theory restricted active space (RASPT2) method for electronic excited states: A benchmark study. *Journal of Chemical Theory and Computation*, 7(1):153–168, 2011. 6.2.1, 6.2.3, 6.2.3

[233] J. Lorentzon, P.-Å. Malmqvist, M. Fülscher, and B.O. Roos. A CASPT2 study of the valence and lowest Rydberg electronic states of benzene and phenol. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, 91(1):91–108, 1995. 6.2.1, 6.2.3

[234] M.R. Silva-Junior, M. Schreiber, S.P.A. Sauer, and W. Thiel. Benchmarks of electronically excited states: Basis set effects on CASPT2 results. *The Journal of Chemical Physics*, 133(17):174318, 2010. 6.2.1, 6.2.3

[235] E.R. Davidson and A.A. Jarzecki. Zero point corrections to vertical excitation energies. *Chemical Physics Letters*, 285(3-4):155–159, 1998. 6.2.1, 6.2.3

[236] N.C. Craig, P. Groner, and D.C. McKean. Equilibrium structures for butadiene and ethylene: Compelling evidence for π-electron delocalization in butadiene. *The Journal of Physical Chemistry A*, 110(23):7461–7469, 2006. 6.2.1, 6.2.3

[237] E.H. Van Veen. Low-energy electron-impact spectroscopy on ethylene. *Chemical Physics Letters*, 41(3):540–543, 1976. 6.2.1

[238] D.F. Evans. Magnetic perturbation of singlet-triplet transitions. Part IV. Unsaturated compounds. *Journal of the Chemical Society*, pages 1735–1745, 1960. 6.2.1, 6.2.3

[239] M. Schreiber, M.R. Silva-Junior, S.P.A. Sauer, and W. Thiel. Benchmarks for electronically excited states: CASPT2, CC2, CCSD, and CC3. *The Journal of Chemical Physics*, 128(13):134110, 2008. 6.2.1, 6.2.3, 6.11, 6.2.3

[240] M.J.S. Dewar and C. De Llano. Ground states of conjugated molecules. XI. Improved treatment of hydrocarbons. *Journal of the American Chemical Society*, 91(4):789–795, 1969. 6.2.2, 6.2.2

[241] E.B. Wilson, J.C. Decius, and P.C. Cross. *Molecular Vibrations*. Dover, New York, 1955. 6.2.2

[242] K. Kuchitsu. *Structure of Free Polyatomic Molecules - Basic Data*. Springer, Berlin/Heidelberg, 1998. 6.2.3

[243] D.R. Lide, editor. *Handbook of Chemistry and Physics*. CRC, Boca Raton, 1997. 6.10, 6.2.3

[244] T.G. Schmalz and L.L. Griffin. How long are the ends of polyene chains? *The Journal of Chemical Physics*, 131(22):224301, 2009. 6.2.3

[245] E.H. Van Veen. Triplet $\pi \rightarrow \pi^*$ transitions in thiophene, furan and pyrrole by low-energy electron-impact spectroscopy. *Chemical Physics Letters*, 41(3):535–539, 1976. 6.2.3

[246] M. Dallos and H. Lischka. A systematic theoretical investigation of the lowest valence- and Rydberg-excited singlet states of trans-butadiene. The character of the 1^1B_u (V) state revisited. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, 112(1):16–26, 2004. 6.2.3

[247] W. Wu, D. Danovich, A. Shurki, and S. Shaik. Using valence bond theory to understand electronic excited states: Application to the hidden excited state (2^1A_g) of $C_{2n}H_{2n+2}$ ($n = 2 - 14$) polyenes. *The Journal of Physical Chemistry A*, 104(38):8744–8758, 2000. 6.2.3

[248] H.A. Witek, H. Nakano, and K. Hirao. Multireference perturbation theory with optimized partitioning. II. Applications to molecular systems. *Journal of Computational Chemistry*, 24(12):1390–1400, 2003. 6.2.3

[249] T. Fang, J. Shen, and S. Li. Block correlated coupled-cluster method with a complete-active-space self-consistent-field reference function: The implementation for low-lying excited states. *The Journal of Chemical Physics*, 129(23):234106, 2008. 6.2.3

[250] B.S. Hudson, B.E. Kohler, and K. Schulten. *Excited States*, volume 6. Academic, New York, 1982. 6.2.3

[251] V.V. Volkov, E.A. Il'inchik, O.V. Volkov, and O.O. Yuryeva. The luminescence of cluster derivatives of boron hydrides, and some applied aspects. *Chemistry for sustainable development*, 8:185–191, 2000. 6.3

[252] Y. Li and L.G. Sneddon. Improved synthetic route to *n*-$B_{18}H_{22}$. *Inorganic Chemistry*, 45(2):470–471, 2006. 6.3.1

[253] M.G.S. Londesborough, D. Hnyk, J. Bould, L. Serrano-Andrés, V. Sauri, J.M. Oliva, P. Kubát, T. Polívka, and K. Lang. Distinct photophysics of the isomers of $B_{18}H_{22}$ explained. *Inorganic Chemistry*, 51(3):1471–1479, 2012. 6.3.1

[254] J.R. Lakowicz. *Principles of Fluorescence Spectroscopy*. Kluwer Academic/Plenum, New York, 1999. 6.3.1

[255] G. Rossbroich, N.A. Garcia, and S.E. Braslavsky. Thermal-lensing measurements of singlet molecular oxygen ($^1\Delta_g$): Quantum yields of formation and lifetimes. *Journal of Photochemistry*, 31(1):37–48, 1985. 6.3.1

[256] P. Å. Malmqvist, B.O. Roos, and B. Schimmelpfennig. The restricted active space (RAS) state interaction approach with spin-orbit coupling. *Chemical Physics Letters*, 357(3-4):230–240, 2002. 6.3.1

[257] G.J. Mains. *Ab initio* molecular orbital study of adducts and oxides of boron hydrides. *The Journal of Physical Chemistry*, 95(13):5089–5096, 1991. 6.3.2

[258] F. Wilkinson, W.P. Helman, and A.B. Ross. Rate constants for the decay and reactions of the lowest electronically excited singlet state of molecular Oxygen in solution. An expanded and revised compilation. *Journal of Physical and Chemical Reference Data*, 24(2):663–677, 1995. 6.3.2

[259] J.R. Platt. Classification of spectra of cata-condensed hydrocarbons. *The Journal of Chemical Physics*, 17(5):484–495, 1949. 6.3.3

[260] K. Lang, J. Mosinger, and D.M. Wagnerová. Photophysical properties of porphyrinoid sensitizers non-covalently bound to host molecules: models for photodynamic therapy. *Coordination Chemistry Reviews*, 248(3-4):321–350, 2004. 6.3.3

[261] J.J. Serrano-Pérez, G. Olaso-González, M. Merchán, and L. Serrano-Andrés. Singlet Oxygen generation in PUVA therapy studied using electronic structure calculations. *Chemical Physics*, 360(1-3):85–96, 2009. 6.3.3

[262] L.M. Frutos and O. Castaño. A new algorithm for predicting triplet-triplet energy-transfer activated complex coordinate in terms of accurate potential-energy surfaces. *The Journal of Chemical Physics*, 123(10):104108, 2005. 6.3.3

[263] L.J. Todd. Recent developments in the study of carboranes. *Pure and Applied Chemistry*, 30(3-4):587–606, 1972. 6.3.3

[264] R. Dahm. Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. *Human Genetics*, 122(6):565–581, 2008. 10.1007/s00439-007-0433-0. 6.4

[265] M.K. Shukla and J. Leszczynski. *Challenges and Advances in Computational Chemistry and Physics*, volume 5, chapter Radiation Induced

Molecular Phenomena in Nucleic Acids: A Comprehensive Theoretical and Experimental Analysis. Springer, The Netherlands, 2008. 6.4

[266] E.C. Friedberg, G.C. Walker, W. Siede, R.D. Wood, R.A. Shultz, and T. Ellenberger. *DNA Repair and Mutagenesis*. ASM Press, Washington, DC, USA, 2006. 6.4

[267] A.O. Colson and M.D. Sevilla. *Computational Molecular Biology*, volume 8, pages 245–270. Elsevier, Amsterdam, 1999. 6.4

[268] J. Leszczynski, editor. *Computational Molecular Biology*. Elsevier, Amsterdam, 1999. 6.4

[269] L.A. Eriksson, editor. *Theoretical Biochemistry. Processes and Properties in Biological Systems*. Elsevier, Amsterdam, 2001. 6.4

[270] L.-Y. Fu, G.-Z. Wang, B.-G. Ma, and H.-Y. Zhang. Exploring the common molecular basis for the universal DNA mutation bias: Revival of Löwdin mutation model. *Biochemical and Biophysical Research Communications*, 409(3):367–371, 2011. 6.4

[271] C.E. Crespo-Hernández, B. Cohen, P. M. Hare, and B. Kohler. Ultrafast excited-state dynamics in nucleic acids. *Chemical Reviews*, 104(4):1977–2020, 2004. 6.4

[272] L. Serrano-Andrés, M. Merchán, and A.C. Borin. A three-state model for the photophysics of adenine. *Chemistry - A European Journal*, 12(25):6559–6571, 2006. 6.4, 6.4.1

[273] L. Serrano-Andrés and M. Merchán. Are the five natural DNA/RNA base monomers a good choice from natural selection?: A photochemical perspective. *Journal of Photochemistry and Photobiology C: Photochemistry Reviews*, 10(1):21–32, 2009. 6.4

[274] M. Merchán and L. Serrano-Andrés. Ultrafast internal conversion of excited cytosine via the lowest $\pi\pi^*$ electronic singlet state. *Journal of the American Chemical Society*, 125(27):8108–8109, 2003. 6.4

[275] M. Merchán, R. González-Luque, T. Climent, L. Serrano-Andrés, E. Rodríguez, M. Reguero, and D. Peláez. Unified model for the ultrafast decay of pyrimidine nucleobases. *The Journal of Physical Chemistry B*, 110(51):26471–26476, 2006. 6.4

[276] L. Serrano-Andrés, M. Merchán, and A.C. Borin. A three-state model for the photophysics of guanine. *Journal of the American Chemical Society*, 130(8):2473–2484, 2008. 6.4, 6.4.1, 6.4.1, 6.4.3, 6.4.3

[277] E. Samoylova, H. Lippert, S. Ullrich, I.V. Hertel, W. Radloff, and T. Schultz. Dynamics of photoinduced processes in adenine and thymine base pairs. *Journal of the American Chemical Society*, 127(6):1782–1786, 2005. 6.4

[278] C. Canuel, M. Mons, F. Piuzzi, B. Tardivel, I. Dimicoli, and M. Elhanine. Excited states dynamics of DNA and RNA bases: Characterization of a stepwise deactivation pathway in the gas phase. *The Journal of Chemical Physics*, 122(7):074316, 2005. 6.4

[279] M.K. Kuimova, J. Dyer, M.W. George, D.C. Grills, J.M. Kelly, P. Matousek, A.W. Parker, X.Z. Sun, M. Towrie, and A.M. Whelan. Monitoring the effect of ultrafast deactivation of the electronic excited states of DNA bases and polynucleotides following 267 nm laser excitation using picosecond time-resolved infrared spectroscopy. *Chemical Communications*, (9):1182–1184, 2005. 6.4

[280] D. Roca-Sanjuaán, G. Olaso-Gonzaález, I. Gonzaález-Ramírez, L. Serrano-Andrés, and M. Merchán. Molecular basis of DNA photodimerization: Intrinsic production of cyclobutane cytosine dimers. *Journal of the American Chemical Society*, 130(32):10768–10779, 2008. 6.4, 6.4.2

[281] I. González-Ramírez, D. Roca-Sanjuán, T. Climent, J. Serrano-Pérez, M. Merchán, and L. Serrano-Andrés. On the photoproduction of DNA/RNA cyclobutane pyrimidine dimers. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, 128(4):705–711, 2011. 6.4, 6.4.2

[282] V.I. Pechenaya, V.I. Danilov, O.N. Slyusarchuk, and J.L. Alderfer. Theoretical investigation of excimer and exciplex states of uracil and halogen derivatives: Effect of nonparallelism of bases. *Photochemistry and Photobiology*, 61(5):435–441, 1995. 6.4

[283] V.I. Danilov, O.N. Slyusarchuk, J.L. Alderfer, J.J.P. Stewart, and P.R. Callis. A theoretical study of the cytosine excimer state: The role of geometry optimization. *Photochemistry and Photobiology*, 59(1):125–129, 1994. 6.4

[284] J. Cadet and P. Vigny. *Bioorganic Photochemistry*, chapter The photochemistry of nucleic acids. John Wiley and Sons, New York, 1990. 6.4

[285] T. Douki and J. Cadet. Individual determination of the yield of the main UV-induced dimeric pyrimidine photoproducts in DNA suggests a high mutagenicity of CC photolesions. *Biochemistry*, 40(8):2495–2501, 2001. 6.4

[286] S. Yamazaki and T. Taketsugu. Photoreaction channels of the Guanine-Cytosine base pair explored by long-range corrected TD-DFT calculations. *Physical Chemistry Chemical Physics*, 14(25):8866–8877, 2012. 6.4

[287] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, and P. A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, 117(19):5179–5197, 1995. 6.4

[288] T.E. Cheatham III, P. Cieplak, and P.A. Kollman. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *Journal of the Biomolecular Structure and Dynamics*, 16(4):845–862, 1999. 6.4

[289] E. Wigner. On the quantum correction for thermodynamic equilibrium. *Physical Review*, 40(5):749–759, 1932. 6.4